

STIC-Biotech/ChemLib

137546

mej

From: Whiteman, Brian  
Sent: Wednesday, November 10, 2004 10:02 AM  
To: STIC-Biotech/ChemLib  
Subject: sequence search

RECEIVED  
NOV 10 2004  
STIC-BIOTECH DIVISION  
(STIC)

10/069,386, Alaoui-Jamali et al.  
2/19/02

search SEQ ID NO: 2 against databases.

Thank you,

Brian Whiteman  
Remsen, 2D14  
mail box 2C18  
Patent Examiner - Art Unit 1635  
United States Patent and Trademark Office  
(571) 272-0764

\*\*\*\*\*

STAFF USE ONLY

Searcher: \_\_\_\_\_  
Searcher Phone: 2-\_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Rev. Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

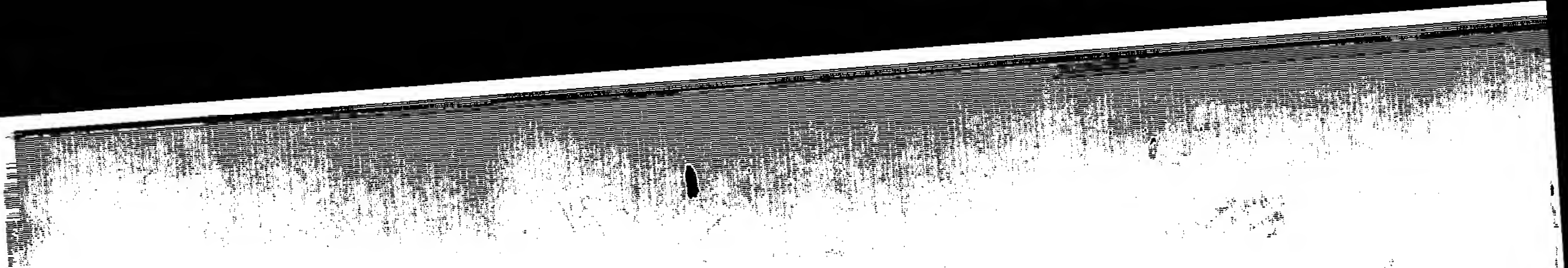
Type of Search

NA Sequence: # \_\_\_\_\_  
AA Sequence: # \_\_\_\_\_  
Structure: # \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other(Specify): \_\_\_\_\_



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2004, 13:49:05 ; Search time 156 Seconds  
(without alignments)  
450.711 Million cell updates/sec

Title: US-10-069-386-2  
Perfect score: 1033  
Sequence: 1 MEGGLKRRKHSLEEEERWE.....APGSWEWNELDHIMEIILGS 196

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score		Query Match %		Length	DB	ID	Description
1	1033	100.0	196	4	AAB35402			Aab35402 Replicati
2	1026	99.3	196	4	AAM93922			Aam93922 Human pol
3	1026	99.3	196	8	ADL32052			Adl32052 Human pro
4	791	76.6	157	4	AAG74742			Aag74742 Human col
5	645.5	62.5	142	8	ABO60075			AbO60075 Human gen
6	402	38.9	111	8	ABO60074			AbO60074 Human gen
7	216	20.9	236	2	AAY36004			Aay36004 Extended
8	216	20.9	236	8	ADP19312			Adp19312 Human sec
9	213	20.6	236	3	AAY44362			Aay44362 Human cel
10	213	20.6	236	4	AAM93724			Aam93724 Human pol
11	213	20.6	236	4	AAG89292			Aag89292 Human sec
12	213	20.6	236	8	ADL31644			Adl31644 Human pro
13	213	20.6	278	4	AAM25550			Aam25550 Human pro
14	211	20.4	236	2	AAY31829			Aay31829 Human adu
15	205	19.8	222	8	ABO60367			AbO60367 Human gen
16	205	19.8	236	2	AAY02619			Aay02619 Amino aci
17	136	13.2	237	5	AAO22897			Aao22897 Mouse hae
18	125	12.1	314	5	ABP65076			Abp65076 Hypoxia-i
19	106	10.3	241	4	AAM39725			Aam39725 Human pol
20	106	10.3	254	3	AAB58258			Aab58258 Lung canc
21	106	10.3	254	4	AAG73682			Aag73682 Human col
22	106	10.3	254	4	AAM41511			Aam41511 Human pol
23	105.5	10.2	578	4	AAM38707			Aam38707 Human pol
24	105.5	10.2	620	7	ADN95361			Adn95361 Human BEC
25	105.5	10.2	650	4	AAM38706			Aam38706 Human pol

26	105.5	10.2	685	7	ADE28201	Ade28201 Human MDD
27	105.5	10.2	759	4	AAM40492	Aam40492 Human pol
28	105.5	10.2	759	4	AAM40493	Aam40493 Human pol
29	105.5	10.2	760	7	ADN95128	Adn95128 Human LEC
30	105	10.2	241	5	ABB06375	Abb06375 Human CHD
31	105	10.2	241	5	AAO22898	Aao22898 Human hae
32	104.5	10.1	740	4	ABB11713	Abb11713 Human KIA
33	104.5	10.1	879	5	ABP69285	Abp69285 Human pol
34	104	10.1	322	4	ABG22868	Abg22868 Novel hum
35	101.5	9.8	728	7	ADD46841	Add46841 Human pro
36	101.5	9.8	728	7	ADE56294	Ade56294 Human pro
37	101.5	9.8	728	8	ADJ66562	Adj66562 PI3 kinas
38	98	9.5	562	7	ADB65060	Adb65060 Human pro
39	96	9.3	817	6	ABR53281	Abr53281 Protein s
40	96	9.3	817	7	ADK63588	Adk63588 Disease t
41	94.5	9.1	728	4	AAM78754	Aam78754 Human pro
42	94.5	9.1	885	8	ADP22490	Adp22490 Sea-squir
43	94	9.1	433	7	ADJ92004	Adj92004 Human T-C
44	94	9.1	743	4	AAM79738	Aam79738 Human pro
45	93.5	9.1	256	8	ADH75872	Adh75872 Corn C-re

ALIGNMENTS

RESULT 1  
AAB35402  
ID AAB35402 standard; protein; 196 AA.  
XX  
AC AAB35402;  
XX  
DT 23-MAY-2001 (first entry)  
XX  
DE Replication protein A binding transcriptional activator 1 RBT1.  
XX  
KW RBT1; replication protein A binding transcriptional activator 1; RPA32;  
KW gene therapy; apoptosis; cancer; leukaemia.  
XX  
OS Unidentified.  
XX  
PN WO200114546-A2.  
XX  
PD 01-MAR-2001.  
XX  
PF 17-AUG-2000; 2000WO-CA000948.  
XX  
PR 19-AUG-1999; 99US-0149472P.  
XX  
(TRAN-) CENT TRANSLATIONAL RES IN CANCER.  
Alaoui-Jamali MA, Cho JM;  
WPI; 2001-218447/22.  
N-PSDB; AAF28052.  
Novel replication protein A binding transcriptional activator 1 gene,  
useful for treating neoplastic disorders such as cancer and in gene  
therapy.  
Disclosure; Fig 1; 16pp; English.  
The present invention provides the protein and coding sequences of the  
replication protein A binding transcriptional activator 1 (RBT1). The  
protein is capable of inducing apoptosis. The sequences are useful in the  
gene therapy and other methods of treatment of cancer, including  
leukaemias. The present sequence is the RBT1 protein

Query Match 100.0%; Score 1033; DB 4; Length 196;  
Best Local Similarity 100.0%; Pred. No. 1.5e-92;  
Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;





Db 1 MVGGLKRKHSDEEEERWESPAGLQSYQQALLRISLDKVQSLGPRAPSLRRHVLHN 60  
QY 61 TLQQLQAALRLAPALPPEPLFLGEEDFSLSATIGSILRELDTSMDGTEPPQNPVTPLG 120  
Db 61 TLQQLQAALRLAPALPPEPLFLGEEDFSLSATIGSILRELDTSMDGTEPPQNPVTPLG 120  
QY 121 LQNEVPPQDPVFLEALSSRYLGDGLDDFFLDIDTSAVEKEPARAPPEPPHNLFCAPGS 180  
Db 121 LQNEVPPQDPVFLEALSSRYLGDGLDDFFLDIDTSAVEKEPARAPPEPPHNLFCAPGS 180  
QY 181 WEWNELDHIMEIILGS 196  
Db 181 WEWNELDHIMEIILGS 196

RESULT 4  
AAG74742  
ID AAG74742 standard; protein; 157 AA.  
XX AAG74742;  
AC  
XX 03-SEP-2001 (first entry)  
DT Human colon cancer antigen protein SEQ ID NO:5506.  
DE  
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KW colorectal carcinoma.  
XX Homo sapiens.  
OS  
XX WO200122920-A2.  
PN  
XX 05-APR-2001.  
PD  
XX 28-SEP-2000; 2000WO-US026524.  
PF  
XX 29-SEP-1999; 99US-0157137P.  
PR 03-NOV-1999; 99US-0163280P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX Ruben SM, Barash SC, Birse CE, Rosen CA;  
PI  
XX WPI; 2001-235357/24.  
DR N-PSDB; AAH34147.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
PT useful for preventing, diagnosing and/or treating colorectal cancers.  
XX  
XX Claim 11; Page 7098-7099; 9803pp; English.  
XX  
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the  
CC proteins are collectively known as colon cancer antigens. The colon  
CC cancer antigens have cytostatic activity and can be used in gene therapy  
CC and vaccine production. N and P may be used in the prevention, diagnosis  
CC and treatment of diseases associated with inappropriate P expression. For  
CC example, N and P may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of P by expressing inactive proteins or to  
CC supplement the patients own production of P. Additionally, N may be used  
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids  
CC into a host cell and culturing the cell to express the proteins. N and P  
CC can be used in the prevention, diagnosis and treatment of colorectal  
CC carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent  
CC sequences used in the exemplification of the present invention. N.B.  
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at  
CC time of publication, meaning no sequences are present for SEQ ID NO:1027  
CC to 1052, 7921 and 7922  
XX  
XX Sequence 157 AA;

Query Match 76.6%; Score 791; DB 4; Length 157;  
Best Local Similarity 98.0%; Pred. No. 5.1e-69;  
Matches 150; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 44 SLGPRAPSLRRHVLHNTLQQLQAALRLAPALPPEPLFLGEEDFSLSATIGSILRELD 103  
Db :|||||  
5 ALGPRAPSLRRHVLHNTLQQLQAALRLAPALPPEPLFLGEEDFSLSATIGSILRELD 64  
QY 104 TSMGTEPPQNPVTPLGLQNEVPPQDPVFLEALSSRYLGDGLDDFFLDIDTSAVEKEP 163  
Db :|||||  
65 TSMGTEPPQNPVTPLGLQNEVPPQDPVFLEALSSRYLGDGLDDFFLDIDTSAVEKEP 124  
QY 164 ARAPPEPPHNLFCAPGSWEWNELDHIMEIILGS 196  
Db :|||||  
125 ARAPPEPPHNLFCAPGSWEWNELDHIMEIILGS 157

RESULT 5  
ABO60075  
ID ABO60075 standard; protein; 142 AA.  
XX  
AC ABO60075;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE Human genome derived single exon protein #6309.  
XX  
KW Human; gene expression; single exon probe; microarray;  
KW alternative splicing event; genomic alteration.  
XX  
OS Homo sapiens.  
XX US2003194704-A1.  
FN  
XX 16-OCT-2003.  
PD  
XX 03-APR-2002; 2002US-00029386.  
PF  
XX 03-APR-2002; 2002US-00029386.  
PR  
XX (PENN/) PENN S G.  
PA (RANK/) RANK D R.  
PA (HANZ/) HANZEL D K.  
XX  
PI Penn SG, Rank DR, Hanzel DK;  
XX  
XX WPI; 2004-119264/12.

XX New human genome-derived single exon nucleic acid probes useful for human  
PT gene expression analysis, for identifying or characterizing alternative  
PT splicing events, for assessing genomic alterations or as tools for  
PT surveying tissues.  
XX  
XX Claim 45; SEQ ID NO 33709; 80pp; English.  
PS  
XX The invention relates to a nucleic acid probe for measuring human gene  
CC expression, comprising any of the 27,400 fully defined nucleotide  
CC sequences in the specification, or their complements or fragments, and  
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences  
CC fully defined in the specification. The probe is a single exon probe that  
CC hybridises under high stringency conditions to a nucleic acid molecule  
CC expressed in human cells or tissues. Also included are a spatially-  
CC addressable set of single exon nucleic acid probes for measuring human  
CC gene expression (comprising a plurality of single exon nucleic acid  
CC probes cited above, where each of the plurality of probes is separately  
CC and addressably isolatable or amplifiable from the plurality), a single  
CC exon microarray for measuring human gene expression, a method of  
CC measuring human gene expression, a vector comprising the single exon  
CC probe cited above, an ORF-encoded peptide comprising at least 8  
CC contiguous amino acids of any of the above-mentioned amino acid  
CC sequences (optionally with conservative amino acid substitutions), an  
CC isolated antibody that binds specifically to a peptide cited above,  
CC methods of selling and/or licensing single exon probes or microarrays to

CC a customer desiring to measure gene expression, a method of providing  
CC human gene expression data by subscription, and a computer-readable  
CC storage medium which contains a database having a plurality of records  
CC (each record including data on the expression of a single exon probe  
CC cited above. The probe, methods and apparatus are useful in gene  
CC expression analysis. The probes may be used as tools for surveying  
CC tissues to detect the presence of expressed messages that contain their  
CC specific exon, or in constructing genome-derived single exon microarrays.  
CC In addition, the probes are used in identifying and characterising  
CC alternative splicing events, in detecting and characterising gross  
CC alterations in the genomic locus that includes their exon, in assessing  
CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
CC or in expressing the ORF-encoded peptide. The present sequence is a human  
CC single exon probe protein of the invention. Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030194704  
XX

Sequence 142 AA;

Query Match 62.5%; Score 645.5; DB 8; Length 142;  
Best Local Similarity 96.8%; Pred. No. 7.4e-55;  
Matches 122; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 71 LAPAPALPPEPLFLGEEDEFSLSATIGSILRELDTSMDGTEPPQNPTPLGLQNEVPPQPD 130  
DB 18 LPPLPC-PPEPLFLGEEDEFSLSATIGSILRELDTSMDGTEPPQNPTPLGLQNEVPPQPD 76  
QY 131 PVFLEALSSRYLGDGLDDFFLDIDTSAVEKEPARAPPEPPHNLFCAPGSWEWNELDHIM 190  
DB 77 PVFLEALSSRYLGDGLDDFFLDIDTSAVEKEPARAPPEPPHNLFCAPGSWEWNELDHIM 136  
QY 191 EIIILGS 196  
DB 137 EIIILGS 142

RESULT 6  
AB060074

ID ABO60074 standard; protein; 111 AA.

XX AB060074;

XX 29-JUL-2004 (first entry)

XX Human genome derived single exon protein #6308.

DE Human; gene expression; single exon probe; microarray;  
XX alternative splicing event; genomic alteration.

XX Homo sapiens.

XX US2003194704-A1.

XX 16-OCT-2003.

XX 03-APR-2002; 2002US-00029386.

XX 03-APR-2002; 2002US-00029386.

XX (PENN/) PENN S G.

XX (RANK/) RANK D R.

XX (HANZ/) HANZEL D K.

XX Penn SG, Rank DR, Hanzel DK;

XX WPI; 2004-119264/12.

XX New human genome-derived single exon nucleic acid probes useful for human  
XX gene expression analysis, for identifying or characterizing alternative  
XX splicing events, for assessing genomic alterations or as tools for  
XX surveying tissues.

PS

XX Claim 45; SEQ ID NO 33708; 80pp; English.

CC The invention relates to a nucleic acid probe for measuring human gene  
CC expression, comprising any of the 27,400 fully defined nucleotide  
CC sequences in the specification, or their complements or fragments, and  
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences  
CC fully defined in the specification. The probe is a single exon probe that  
CC hybridises under high stringency conditions to a nucleic acid molecule  
CC expressed in human cells or tissues. Also included are a spatially-  
CC addressable set of single exon nucleic acid probes for measuring human  
CC gene expression (comprising a plurality of single exon nucleic acid  
CC probes cited above, where each of the plurality of probes is separately  
CC and addressably isolatable or amplifiable from the plurality), a single  
CC exon microarray for measuring human gene expression, a method of  
CC measuring human gene expression, a vector comprising the single exon  
CC probe cited above, an ORF-encoded peptide comprising at least 8  
CC contiguous amino acids of any of the above-mentioned amino acid  
CC sequences (optionally with conservative amino acid substitutions), an  
CC isolated antibody that binds specifically to a peptide cited above,  
CC methods of selling and/or licensing single exon probes or microarrays to  
CC a customer desiring to measure gene expression, a method of providing  
CC human gene expression data by subscription, and a computer-readable  
CC storage medium which contains a database having a plurality of records  
CC (each record including data on the expression of a single exon probe  
CC cited above. The probe, methods and apparatus are useful in gene  
CC expression analysis. The probes may be used as tools for surveying  
CC tissues to detect the presence of expressed messages that contain their  
CC specific exon, or in constructing genome-derived single exon microarrays.  
CC In addition, the probes are used in identifying and characterising  
CC alternative splicing events, in detecting and characterising gross  
CC alterations in the genomic locus that includes their exon, in assessing  
CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
CC or in expressing the ORF-encoded peptide. The present sequence is a human  
CC single exon probe protein of the invention. Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030194704  
XX

SQ Sequence 111 AA;

Query Match 38.9%; Score 402; DB 8; Length 111;  
Best Local Similarity 89.9%; Pred. No. 3.4e-31;  
Matches 80; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 MEGGLKRKHSDEEEERWEWSPAGLQSYQALLRISLDKVKQSLGPRAPSLRRHVLIHN 60  
DB 3 MVGGLKRKHSDEEEERWEWSPAGLQSYQALLRISLDKVKQSLGPRAPSLRRHVLIHN 62

QY 61 TLQQLQAALRLAPALPPEPLFLGEEEDF 89

DB 63 TLQQLQAALRLAPALPPEPLFLGEEEDF 91

RESULT 7

AAAY36004

ID AAY36004 standard; protein; 236 AA.

XX AAY36004;

XX 13-SEP-1999 (first entry)

XX Extended human secreted protein sequence, SEQ ID NO. 389.

XX Secreted protein; human; cytokine; cellular proliferation; cell movement;  
XX cellular differentiation; immune system regulator; anti-inflammatory;  
XX haematopoiesis regulator; tissue growth regulator; tumour inhibitor;  
XX reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;  
XX genetic disease.

XX Homo sapiens.

XX WO9931236-A2.

```
PD 24-JUN-1999.
XX
PF 17-DEC-1998; 98WO-IB002122.
XX
PR 17-DEC-1997; 97US-0069957P.
PR 09-FEB-1998; 98US-0074121P.
PR 13-APR-1998; 98US-0081563P.
PR 10-AUG-1998; 98US-0096116P.
XX
PA (GEST ) GENSET.
XX
PI Bougueleret L, Duclert A, Dumas Milne Edwards J;
XX
DR WPI; 1999-385906/32.
DR N-PSDB; AAX97688.
XX
PT New isolated human secreted proteins.
XX
PS Claim 9; Page 330-331; 516pp; English.
XX
CC This sequence is encoded by an extended human secreted protein coding
CC sequence of the invention. The secreted proteins can be used in treating
CC or controlling a variety of human conditions. The secreted proteins may
CC act as cytokines or may affect cellular proliferation or differentiation
CC or may act as immune system regulators, haematopoiesis regulators, tissue
CC growth regulators, regulators of reproductive hormones or cell movement
CC or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
CC tumour inhibition activity. The DNAs can be used in forensic procedures
CC to identify individuals or in diagnostic procedures to identify
CC individuals having genetic diseases resulting from abnormal expression of
CC the genes corresponding to the extended cDNAs. They are also useful for
CC constructing a high resolution map of the human chromosomes. They can
CC also be used for gene therapy to control or treat genetic diseases
XX
SQ Sequence 236 AA;
Query Match 20.9%; Score 216; DB 2; Length 236;
Best Local Similarity 32.6%; Pred. No. 1.3e-12;
Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;
QY 4 GLKRKHSDEEEE---ERWESPAGLQSYQQA-----LLRLSLDKVQRSLGPPRAPSS 51
Db 5 GLKKRKEEEEEKPEPLAVDSW-WLDPGHAAVAQAAPPVASSSFLDLSVLKLHHSLQXSXP 63
QY 52 LRRHVLIHTNQLOALRALAPALPPEPL-----FLGEEDFSLSATIGSILRE 101
Db 64 LRHLVLXNTLRRIQAS--MAPAAALPVPTTPAAPXVADNLLASSDAALSASMAXLLED 121
QY 102 LDTSMDGTETPNVTPLGLQNEVPQPDP----PVFLEAL-----SSRYLGSDGLDDFFL 152
Db 122 L-SHIEGLSQAPQP-----LADEGPGRSIGGXPPXLGALLDLGGPATGCCLLDNGLEG 175
QY 153 DIDTSAVEKE---PARAPPEP-PHNLFCAFGSWE-----WNELDHIMEIILGS 196
Db 176 DIDTSMYDNLWAPASEGLKPGED---GPKKEEAPELDEALDYLMVDLVGT 225
RESULT 8
ADP19312
ID ADP19312 standard; protein; 236 AA.
XX AC ADP19312;
XX DT 26-AUG-2004 (first entry)
XX DE Human secreted polypeptide #163.
XX KW Human; secreted protein; genetic disease.
XX OS Homo sapiens.
XX PN US2004110939-A1.
XX
```









64 LRHLVLVNTLRRIQAS--MAPAAALPPVPSPPAAPSVADNLLASSDAALSASMASLLED 121  
102 LDTSMGTEPPQNPTPLGLQNEVPP-----QDPVFLEAL--SSRYLGDSDLDDFFL 152  
122 L-SHIEGLSQAPQP-----LADEGPPGRSIGGAAPSLGALDILGPATGCLDDGLEGLFE 175  
153 DIDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEIILGS 196  
176 DIDTSMYDNLWAPASEGLKPGPED---GPGKEEAPELDEALDYLMVDVLVGT 225

RESULT 13  
AAM25550  
ID AAM25550 standard; protein; 278 AA.

XX AAM25550;

DT 16-OCT-2001 (first entry)

DE Human protein sequence SEQ ID NO:1065.

XX Human; cancer; ulcer; HIV infection; human immunodeficiency virus;  
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;  
KW antibacterial; endocrine; cardiant; central nervous system; virucide;  
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;  
KW antiaggregant; haemostatic; vulnery; antiulcer; osteopathic; cytostatic;  
KW dermatological; antiallergic; antiasthmatic; antidiabetic; infection;  
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; inflammation;  
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;  
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;  
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;  
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;  
KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;  
KW allergic rhinitis; diabetes; multiple sclerosis; depression;  
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
KW neurological disorder.

XX Homo sapiens.

XX WO200153455-A2.

XX 26-JUL-2001.

DE 22-DEC-2000; 2000WO-US035017.

PR 23-DEC-1999; 99US-00471275.

PR 21-JAN-2000; 2000US-00488725.

PR 25-APR-2000; 2000US-00552317.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-457603/49.

DR N-PSDB; AAH99491.

XX Isolated human polynucleotides encoding polypeptides, useful for the

XX treatment and diagnosis of e.g. cancer, ulcers and HIV infection.

PS Claim 20; Page 214; 1217pp; English.

XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to

CC AAM25963. The proteins can have activities based on the tissues and cells

CC they are expressed in, such as: antiinflammatory; antirheumatic;

CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;

CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;

CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;

CC antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;

CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;

CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides

CC encoding them can be used in gene therapy, antisense therapy and vaccine

CC production. The proteins and polynucleotides are useful for screening for

CC agonists or antagonists of a protein and for the treatment and diagnosis

CC of disorders associated with the activity of a protein e.g. inflammation,  
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,  
CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,  
CC osteoporosis, severe combined immunodeficiency, eczema, allergic  
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and  
CC neurological disorders

XX Sequence 278 AA;

Query Match 20.6%; Score 213; DB 4; Length 278;

Best Local Similarity 32.6%; Pred. No. 3.2e-12;

Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

QY 4 GLKRKHSLEEEE---ERWEWSAGLQSYQQA-----LLRISLDKVQSRSLGPRAPS 51

Db 47 GLKRKREEEEEEKEPLAVDSW-WLDPGHAAVAQAAPPAVASSSLFDLSVLKHHSLQQSEPD 105

QY 52 LRRHVLHNTLQQLQAALRLAPALPEPL-----FLGEEDFSLSATIGSILRE 101

Db 106 LRHLVLVNTLRRIQAS--MAPAAALPPVPSPPAAPSVADNLLASSDAALSASMASLLED 163

QY 102 LDTSMGTEPPQNPTPLGLQNEVPP-----QDPVFLEAL--SSRYLGDSDLDDFFL 152

Db 164 L-SHIEGLSQAPQP-----LADEGPPGRSIGGAAPSLGALDILGPATGCLDDGLEGLFE 217

QY 153 DIDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEIILGS 196

Db 218 DIDTSMYDNLWAPASEGLKPGPED---GPGKEEAPELDEALDYLMVDVLVGT 267

RESULT 14

AAAY31829

ID AAY31829 standard; protein; 236 AA.

XX AAY31829;

XX 06-DEC-1999 (first entry)

DT Human adult blood secreted protein g21\_1.

DE Secreted protein; g21\_1; human; therapy; diagnosis; vaccine; blood.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "encoded by AWG"

FT Peptide 19..21 /note= "signal peptide"

FT Protein 32..236 /note= "mature protein"

FT Domain 40 /note= "a putative transmembrane domain is centered

FT Peptide around this residue"

FT Protein 67..79 /note= "alternative signal peptide"

FT Domain 80..236 /note= "alternative mature protein"

FT Misc-difference 137 /note= "a putative transmembrane domain is centered

FT Protein around this residue"

FT Domain 150 /note= "encoded by GAS"

FT Misc-difference 150 /note= "a putative transmembrane domain is centered

FT Protein around this residue"

XX WO9947555-A1.

XX 23-SEP-1999.

PD



us-10-069-386-2.rag

Mon Nov 15 17:33:06 2004

122 L-SHIEGLSQAPQP-----LADEGPPGRSIGGAAPSLGALDOLLGPATGCLLDDGLEGLFE 175  
153 DIDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEII 193  
176 DIDTSMYDNELWAPASEGLKPGPED---GPGKEEAPELDEAELDYLMVDL 222

Search completed: November 15, 2004, 14:03:14  
Job time : 159 secs





GN Name=SERTAD3;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Ovary, and Uterus;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Ovary;  
RA Strausberg R.;  
RX Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Uterus;  
RA Strausberg R.;  
RX Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.  
RL EMBL; BC014061; AAH14061.1; -;  
DR EMBL; BC050643; AAH50643.1; -;  
DR InterPro; IPR009263; SERTA.  
DR Pfam; PF06031; SERTA; 1.  
SQ SEQUENCE 196 AA; 21768 MW; 9C54AFA126F17AF1 CRC64;  
  
Query Match 99.3%; Score 1026; DB 2; Length 196;  
Best Local Similarity 99.5%; Pred. No. 3.9e-72;  
Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 MEGGLKRXHSDLEEEERWESPAGLQSYQQALLRISLDKVQVSLGPRAPSLRRHVLIHN 60  
DB 1 MVGGLKRXHSDLEEEERWESPAGLQSYQQALLRISLDKVQVSLGPRAPSLRRHVLIHN 60  
  
QY 61 TLQQLQAALRLAPALPPEPLFLGEEDEFSLSATIGSILRELDTSMDGTEPPQNPVPLG 120  
DB 61 TLQQLQAALRLAPALPPEPLFLGEEDEFSLSATIGSILRELDTSMDGTEPPQNPVPLG 120  
  
QY 121 LQNEVPPQDPVFLEALSSRYLGDGLDDFFLDIDTSAVEKEPARAPPEPPHNLFCAPGS 180  
DB 121 LQNEVPPQDPVFLEALSSRYLGDGLDDFFLDIDTSAVEKEPARAPPEPPHNLFCAPGS 180  
  
QY 181 WEWNELDHIMEIILGS 196  
DB 181 WEWNELDHIMEIILGS 196  
  
RESULT 3  
Q9ERC3 PRELIMINARY; PRT; 197 AA.  
AC Q9ERC3;  
DE 01-MAR-2001 (TrEMBLrel. 16, Created)  
DI 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Replication protein-binding trans-activator RBT1 (SERTA domain  
containing 3).  
GN Name=Sertad3;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RA Strausberg R.;  
RX Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.  
RL EMBL; AF317202; AAG30951.1; -;  
DR EMBL; BC034886; AAH34886.1; -;  
DR MGD; MGI:2180697; Sertad3.  
DR InterPro; IPR000194; ATPase\_a/bcentre.  
DR InterPro; IPR009263; SERTA.  
DR Pfam; PF06031; SERTA; 1.  
DR PROSITE; PS00152; ATPASE ALPHA BETA; UNKNOWN 1.  
SQ SEQUENCE 197 AA; 21934 MW; CFF6FA4C0A91D7E7 CRC64;  
  
Query Match 84.3%; Score 871; DB 2; Length 197;  
Best Local Similarity 85.7%; Pred. No. 4.6e-60;  
Matches 169; Conservative 8; Mismatches 18; Indels 2; Gaps 1;  
  
QY 3 GGLKRXHSDL--EEEEERWESPAGLQSYQQALLRISLDKVQVSLGPRAPSLRRHVLIHN 60  
DB 2 GGLKRXHSDLEEEEEKWDWSPALTALRSYQQALLRISLDKVQVSLGPRAPSLRRHVLIHN 61  
  
QY 61 TLQQLQAALRLAPALPPEPLFLGEEDEFSLSATIGSILRELDTSMDGTEPPQNPVPLG 120  
DB 62 TLQQLQAALRLAPALPPEPLFLGEEDEFSLSTTIGSILRELDTSMDGTEPPQNPVPLG 121  
  
QY 121 LQNEVPPQDPVFLEALSSRYLGDGLDDFFLDIDTSAVEKEPARAPPEPPHNLFCAPGS 180  
DB 122 PQNEIVSQADPVFLEALSSRYLGDGLDDFFLDIDTSAVEKVALPPPEPPHSLFCSPGS 181  
  
QY 181 WEWNELDHIMEIILGS 196  
DB 182 WEWNELDHIMEIILGS 197  
  
RESULT 4  
STD1\_MOUSE  
ID STD1\_MOUSE STANDARD; PRT; 236 AA.











```

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi P., Tanaka T.,
RA Teijima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK002324; BAB22013.1; -
DR MGD; MGI:1913438; Sertad1.
DR InterPro; IPR009263; SERTA.
DR Pfam; PF06031; SERTA; 1.
DR SQA SEQUENCE 244 AA; 26068 MW; 0F464F6419D362A6 CRC64;
Query Match 16.1%; Score 166; DB 2; Length 244;
Best Local Similarity 25.9%; Pred. No. 4.8e-05;
Matches 67; Conservative 33; Mismatches 55; Indels 104; Gaps 13;
OY 4 GLKRRKHSOLEEEERWE-----W-----SPAGLQS---YQALLRISLDKQVRSIGPR 48
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 5 GLKRN-----GEEETMEALSVDSCLDPSPHAPVAQTPTVASSLFDLSVVKLHSLRQS 60
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 49 APSLRRHVLHNTLQQLQAALRLAPALPEPL-----PLGEEDFSLSATIGSI 98
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 61 EPDLRLHLVLVNTLRLRIQASME--PAPVLPPEPIQPAPSPVADSLASSDAGLSASMASL 118
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 99 LRELDTSMDGTEPPQ-----NPVTP-----LG-----LQNEV 125
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 119 LEDLNHIEDLNQAPQPAQDEGPPGRSIGGISPNLGGALDLGQPLAVCWMTMDWRACLRISI 178
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 126 PE-----QPDPVFEALSSRYLGDSDLDDFFLDITSAVEKEPARAPPEPPHNLFCFA 177
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 179 PFCTTVNYGYQP-----LRVSSAAPENGPAK--EEPP----- 208
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 178 PGSEWNEWELDHIMEIILGS 196
DB : || : || : || : || : || : || : || : || : || : || : || : || : ||
OY 209 --ELDEAEELTDLMDVLVGT 225
DB : || : || : || : || : || : || : || : || : || : || : || : || : ||

RESULT 9
STD2_MOUSE STANDARD; PRT; 309 AA.
AC Q9JUG5; Q8C609; Q91WL3; Q925E5;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE SERTA domain-containing protein 2 (Transcriptional regulator
DE interacting with the PHD-bromodomain 2) (TRIP-Br2).
GN Name=Sertad2; Synonyms=Kiaa0127;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
```





GN Name=Cdca4; Synonyms=HEPP;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Bone marrow;  
RX MEDLINE=21375891; PubMed=11482882; DOI=10.1006/bcmd.2001.0434;  
RA Abqilah J.M., Jing X., Spassov D.S., Nachtman R.G., Jurecic R.;  
RT "Cloning and characterization of Hepp, a novel gene expressed  
preferentially in hematopoietic progenitors and mature blood cells."  
RL Blood Cells Mol. Dis. 27:667-676(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Mesonephros;  
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,  
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
60,770 full-length cDNAs."  
RL Nature 420:563-573(2002).  
RN [3]  
RP SEQUENCE OF 13-237 FROM N.A.  
RC TISSUE=Breast tumor;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raba S.S., Logucliano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -!- FUNCTION: May be involved in molecular regulation of hematopoietic  
stem cells and progenitor cell lineage commitment and  
differentiation.

CC -!- TISSUE SPECIFICITY: Expressed preferentially in hematopoietic  
progenitors and mature blood cells. Expressed at low levels in the  
heart, lung, spleen, and thymus and at a higher level in muscle.  
CC -!- DEVELOPMENTAL STAGE: Developmentally regulated. Preferential  
expression in both fetal and adult hematopoietic progenitors and  
mature blood cells during embryonic and adult hematopoiesis.  
CC -!- SIMILARITY: Belongs to the TRIP-Br family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
at the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AK010535; BAB27012.2; --  
CC EMBL; AF322238; AAK31074.1; --  
CC EMBL; AK032980; BAC28109.1; --  
CC EMBL; BC012953; AAH12953.1; ALT\_INIT.  
CC MGD; MGI:1919213; Cdca4.  
CC InterPro; IPR009263; SERTA.  
CC Pfam; PF06031; SERTA; 1.  
CC CONFLICT 196 196 S -> T (in Ref. 3).  
CC CONFLICT 213 213 T -> A (in Ref. 3).  
CC SEQUENCE 237 AA; 26107 MW; 52EC046EFD326E CRC64;  
Query Match 13.2%; Score 136; DB 1; Length 237;  
Best Local Similarity 26.1%; Pred. No. 0.01;  
Matches 63; Conservative 32; Mismatches 86; Indels 60; Gaps 9;  
QY 4 GLKRKHSDLEREREWSPA-GLQSYQQALLRISLDKQVRSGLPRAPSLRRHVLHNTL 62  
DB 5 GLKRKYGDQEGVEGFTVPSYSLQ--PQLDMSLVKLQCHMLVEPNLCRSVLIAITV 62  
QY 63 QQLQAALR-----LAP--APALPPEPLFGEEDFSLSATIGSLRELDTSMDGTEPPQ 113  
DB 63 RQIQEEMSQDGVHGMAPQNVDRAPVERLVSTE---ILCRTVGAEEHPAPELEDAPLQ 119  
QY 114 NPVTPLGLQNEVPPQDP-----VFLEALSSRYLGDSDLDDFF 151  
DB 120 NSVSELPIVGSAPQORNPQSSLWEMDSPQENRGSFQKSLDQIFETLENK--NSSSVEELF 177  
QY 152 LDIDTSAVEKEBPAP-----APPEPHNLFCAQSGWENELDHIMEII 193  
DB 178 SDVDSYDLDTLVTGMSGTSKSLCNCLEGFAAATPPPSSTCKS---DLAELDHVVEIL 234  
QY 194 L 194  
DB 235 V 235  
RESULT 12  
AAH55824  
ID AAH55824 PRELIMINARY; PRT; 237 AA.  
AC AAH55824;  
DT 02-MAR-2004 (TREMBlrel. 27, Created)  
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)  
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)  
DE Cell division cycle associated 4.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Eye;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raba S.S., Logucliano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -!- FUNCTION: May be involved in molecular regulation of hematopoietic  
stem cells and progenitor cell lineage commitment and  
differentiation.







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2004, 14:06:47 ; Search time 142 Seconds  
(without alignments)  
488.370 Million cell updates/sec

Title: US-10-069-386-2  
Perfect score: 1033  
Sequence: 1 MEGGLKRKHSDEEEERWE.....APGSWEWNELDHIMEILGS 196

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1568699 seqs, 353819137 residues

Total number of hits satisfying chosen parameters: 1568699

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18:	/cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	791	76.6	157	14	US-10-106-698-5516 Sequence 5516, Ap
2	645.5	62.5	142	14	US-10-029-386-33709 Sequence 33709, A
3	402	38.9	111	14	US-10-029-386-33708 Sequence 33708, A
4	216	20.9	236	11	US-09-978-360A-568 Sequence 568, App
5	213	20.6	236	9	US-09-731-872-412 Sequence 412, App
6	213	20.6	236	10	US-09-876-997-412 Sequence 412, App
7	213	20.6	278	15	US-10-296-115-1065 Sequence 1065, Ap
8	205	19.8	222	14	US-10-029-386-34001 Sequence 34001, A
9	136	13.2	237	13	US-10-076-069-2 Sequence 2, Appli
10	125	12.1	314	14	US-10-170-385-31 Sequence 31, Appl
11	125	12.1	314	16	US-10-755-889-627 Sequence 627, App
12	106	10.3	254	9	US-09-925-302-596 Sequence 596, App
13	106	10.3	254	10	US-09-925-302-596 Sequence 596, App

14	106	10.3	254	14	US-10-106-698-4456	Sequence 4456, Ap
15	105	10.2	241	13	US-10-076-069-4	Sequence 4, Appli
16	104.5	10.1	740	15	US-10-276-774-2083	Sequence 2083, Ap
17	98	9.5	562	14	US-10-104-047-3214	Sequence 3214, Ap
18	97	9.4	803	14	US-10-369-493-2536	Sequence 2536, Ap
19	97	9.4	1099	17	US-10-425-115-195088	Sequence 195088, A
20	96	9.3	297	15	US-10-425-114-64002	Sequence 64002, A
21	96	9.3	817	14	US-10-369-493-1813	Sequence 1813, Ap
22	95.5	9.2	1228	16	US-10-437-963-188300	Sequence 188300, A
23	95	9.2	278	17	US-10-739-930-8565	Sequence 8565, Ap
24	95	9.2	713	16	US-10-437-963-137248	Sequence 137248, A
25	94	9.1	433	8	US-08-259-451-3	Sequence 3, Appli
26	94	9.1	433	14	US-10-224-999A-3472	Sequence 3472, Ap
27	93.5	9.1	256	14	US-10-421-138A-312	Sequence 312, App
28	93.5	9.1	256	15	US-10-374-780A-1247	Sequence 1247, Ap
29	93.5	9.1	663	16	US-10-755-889-480	Sequence 480, App
30	93.5	9.1	1734	9	US-09-862-027-81	Sequence 81, Appl
31	93.5	9.1	1734	15	US-10-042-865-82	Sequence 82, Appl
32	93	9.0	157	16	US-10-767-701-54984	Sequence 54984, A
33	92.5	9.0	113	13	US-10-076-069-6	Sequence 6, Appli
34	92	8.9	1103	16	US-10-437-963-166325	Sequence 166325, A
35	91.5	8.9	714	15	US-10-425-114-62753	Sequence 62753, A
36	91	8.8	316	15	US-10-425-114-56782	Sequence 56782, A
37	90.5	8.8	690	17	US-10-425-115-221982	Sequence 221982, A
38	90.5	8.8	206	16	US-10-437-963-187376	Sequence 187376, A
39	90.5	8.8	356	17	US-10-425-115-257033	Sequence 257033, A
40	90.5	8.8	370	14	US-10-192-381-16	Sequence 16, Appl
41	90.5	8.8	824	14	US-10-226-844-1	Sequence 1, Appli
42	90.5	8.8	824	14	US-10-210-951-58	Sequence 58, Appl
43	90.5	8.8	824	14	US-10-211-884-58	Sequence 58, Appl
44	90.5	8.8	824	14	US-10-211-858-58	Sequence 58, Appl
45	90.5	8.8	947	9	US-09-871-889-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-10-106-698-5516, Application US/10106698  
; Sequence 5516, Application US/10106698  
; Publication No. US20030109690A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide  
; FILE REFERENCE: PA005P1  
; CURRENT APPLICATION NUMBER: US/10/106,698  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: PCT/US00/26524  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 60/157,137  
; PRIOR FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: US 60/163,280  
; PRIOR FILING DATE: 1999-11-03  
; NUMBER OF SEQ ID NOS: 8564  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 5516  
; LENGTH: 157  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (55)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: MISC FEATURE  
; LOCATION: (132)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-106-698-5516

Query Match 76.6%; Score 791; DB 14; Length 157;  
Best Local Similarity 98.0%; Pred. No. 1.3e-63;  
Matches 150; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 44 SLGPRAPSLRRHVLIHNTLQQLQAALRLAPAPALPPEPLFGLGEEDFSLSATIGSILRELD 103

Db 5 ALGPRAPSLRRHVLIHNTLQQLQAALRLAPALPPEPLFLGDEEDFSLXIGSILRELD 64

Qy 104 TSMGTEPPQNPVTPLGLQNEVPPQDPVFLEALSSRYLGDGSLDDFFLDIDTSAVEKEP 163

Db 65 TSMGTEPPQNPVTPLGLQNEVPPQDPVFLEALSSRYLGDGSLDDFFLDIDTSAVEKEP 124

Qy 164 ARAPPEPHNLFCAPGSWENWELDHIMEILGS 196

Db 125 ARAPPEPHNLFCAPGSWENWELDHIMEILGS 157

RESULT 2

US-10-029-386-33709

Sequence 33709, Application US/10029386

Publication No. US20030194704A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C

TITLE OF INVENTION: EXPRESSION ANALYSIS TWO

FILE REFERENCE: AEOMICA-X-2

CURRENT APPLICATION NUMBER: US/10/029,386

CURRENT FILING DATE: 2001-12-20

NUMBER OF SEQ ID NOS: 34298

SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

SEQ ID NO 33709

LENGTH: 142

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AC010271.5

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.64

US-10-029-386-33709

Query Match 62.5%; Score 645.5; DB 14; Length 142;

Best Local Similarity 96.8%; Pred. No. 1.7e-50;

Matches 122; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 71 LAPAPALPPEPLFLGDEEDFSLSATIGSILRELDTSMDGTEPPQNPVTPLGLQNEVPPQPD 130

Db 18 LPPLPC-PPEPLFLGDEEDFSLSATIGSILRELDTSMDGTEPPQNPVTPLGLQNEVPPQPD 76

Qy 131 PVFLEALSSRYLGDGSLDDFFLDIDTSAVEKEPARAPPEPHNLFCAPGSWENWELDHIM 190

Db 77 PVFLEALSSRYLGDGSLDDFFLDIDTSAVEKEPARAPPEPHNLFCAPGSWENWELDHIM 136

Qy 191 EIILGS 196

Db 137 EIILGS 142

RESULT 3

US-10-029-386-33708

Sequence 33708, Application US/10029386

Publication No. US20030194704A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C

TITLE OF INVENTION: EXPRESSION ANALYSIS TWO

FILE REFERENCE: AEOMICA-X-2

CURRENT APPLICATION NUMBER: US/10/029,386

CURRENT FILING DATE: 2001-12-20

NUMBER OF SEQ ID NOS: 34288

SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

SEQ ID NO 33708

LENGTH: 111

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AC010271.5

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.64

US-10-029-386-33708

Query Match 38.9%; Score 402; DB 14; Length 111;

Best Local Similarity 89.9%; Pred. No. 1.4e-28;

Matches 80; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 MEGGLKRXKHSDEEEERWESPAGLQSYQQALLRISLDKVKVQSLGPRAPSLRRHVLIHN 60

Db 3 MVGGLKRXKHSDEEEERWESPAGLQSYQQALLRISLDKVKVQSLGPRAPSLRRHVLIHN 62

Qy 61 TLQQLQAALRLAPALPPEPLFLGDEEDF 89

Db 63 TLQQLQAALRLAPALPPEPLPGRGGF 91

RESULT 4

US-09-978-360A-568

Sequence 568, Application US/09978360A

Publication No. US20040110939A1

GENERAL INFORMATION:

APPLICANT: Edwards, Jean-Baptiste Dumas Milne

APPLICANT: Duclert, Aymeric

APPLICANT: Bougueleret, Lydie

APPLICANT: Jobert, Severin

APPLICANT: Clusel, Catherine

TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides

FILE REFERENCE: 56.US4.CIP

CURRENT APPLICATION NUMBER: US/09/978,360A

CURRENT FILING DATE: 2001-10-15

PRIOR APPLICATION NUMBER: US 60/066,677

PRIOR FILING DATE: 1997-11-13

PRIOR APPLICATION NUMBER: US 60/069,957

PRIOR FILING DATE: 1997-12-17

PRIOR APPLICATION NUMBER: US 60/074,121

PRIOR FILING DATE: 1998-02-09

PRIOR APPLICATION NUMBER: US 60/081,563

PRIOR FILING DATE: 1998-04-13

PRIOR APPLICATION NUMBER: US 60/096,116

PRIOR FILING DATE: 1998-08-10

PRIOR APPLICATION NUMBER: US 60/099,273

PRIOR FILING DATE: -09-04

PRIOR APPLICATION NUMBER: US 09/191,997

PRIOR FILING DATE: 1998-11-13

PRIOR APPLICATION NUMBER: US 09/215,435

PRIOR FILING DATE: 1998-12-17

PRIOR APPLICATION NUMBER: PCT/IB98/02122

PRIOR FILING DATE: 1998-12-17

PRIOR APPLICATION NUMBER: US 09/247,155

PRIOR FILING DATE: 1999-02-09

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 810

SOFTWARE: Patent.pm

SEQ ID NO 568

LENGTH: 236

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SIGNAL

LOCATION: -31..-1

FEATURE:

NAME/KEY: misc feature

LOCATION: (28, 30, 40, 67, 86, 117, 120)

OTHER INFORMATION: unknown

US-09-978-360A-568

Query Match 20.9%; Score 216; DB 11; Length 236;

Best Local Similarity 32.6%; Pred. No. 2.4e-11;

Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

Qy 4 GLKRXKHSDEEE-----ERWESPAGLQSYQQA-----LLRISLDKVKVQSLGPRAPS 51



Db 5 GLKRKREEEKEPLAVDSW-WLDPGHAAVAQAPPAVASSSLFDLSVLKLHHSLOXSXP 63

QY 52 LRRHVLHNTLQQLQAALRLAPALPPEPL-----FLGEEDFSLSATIGSILRE 101

Db 64 LRHLVLVXNTRLRIQAS--MAPAAALPPVTPPAAPXVADNLLASSDAALSASMAXLLED 121

QY 102 LDTSMGTEPPQNPTPLGLQNEVPPD-----PVFLEAL-----SSRYLGDSGLDDFFL 152

Db 122 L-SHIEGLSQAPQ-----LADEGPPGRSIGGAPXPLGALDLLGPATGCLLDNGLEGLFE 175

QY 153 DIDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEIILGS 196

Db 176 DIDTSMYDNELWAPASEGLKPGPED---GPGKEEAPELDEAELDYLMDVLVGT 225

RESULT 5

US-09-731-872-412

; Sequence 412, Application US/09731872

; Patent No. US20020102604A1

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, Jean Baptiste

; APPLICANT: Bougueleret, Lydie

; APPLICANT: Jobert, Severin

; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS

; FILE REFERENCE: 78.US3.REG

; CURRENT APPLICATION NUMBER: US/09/731,872

; CURRENT FILING DATE: 2000-12-07

; PRIOR APPLICATION NUMBER: US 60/169,629

; PRIOR FILING DATE: 1999-12-08

; PRIOR APPLICATION NUMBER: US 60/187,470

; PRIOR FILING DATE: 2000-03-06

; NUMBER OF SEQ ID NOS: 482

; SOFTWARE: Patent.pm

; SEQ ID NO 412

; LENGTH: 236

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-731-872-412

Query Match 20.6%; Score 213; DB 9; Length 236;

Best Local Similarity 32.6%; Pred. No. 4.6e-11;

Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

QY 4 GLKRKHSDELEEE---ERWEWSAGLQSYQQA-----LLRISLDKVQSRSLGPRAPS 51

Db 5 GLKRKREEEKEPLAVDSW-WLDPGHTAVAQAPPAVASSSLFDLSVLKLHHSLOQSEPD 63

QY 52 LRRHVLHNTLQQLQAALRLAPALPPEPL-----FLGEEDFSLSATIGSILRE 101

Db 64 LRHLVLVXNTRLRIQAS--MAPAAALPPVTPPAAPXVADNLLASSDAALSASMAXLLED 121

QY 102 LDTSMGTEPPQNPTPLGLQNEVPPD-----QPDVFLEAL--SSRYLGDSGLDDFFL 152

Db 122 L-SHIEGLSQAPQ-----LADEGPPGRSIGGAAPSLGALDLLGPATGCLLDNGLEGLFE 175

QY 153 DIDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEIILGS 196

Db 176 DIDTSMYDNELWAPASEGLKPGPED---GPGKEEAPELDEAELDYLMDVLVGT 225

RESULT 6

US-09-876-997-412

; Sequence 412, Application US/09876997

; Publication No. US20030152921A1

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, Jean Baptiste

; APPLICANT: Bougueleret, Lydie

; APPLICANT: Jobert, Severin

; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS

; FILE REFERENCE: 78.US4.CIP

; CURRENT APPLICATION NUMBER: US/09/876,997

; CURRENT FILING DATE: 2001-06-08

; PRIOR APPLICATION NUMBER: US 09/731,872

; PRIOR FILING DATE: 2000-12-07

; PRIOR APPLICATION NUMBER: US 60/187,470

; PRIOR FILING DATE: 2000-03-06

; PRIOR APPLICATION NUMBER: US 60/169,629

; PRIOR FILING DATE: 1999-12-08

; NUMBER OF SEQ ID NOS: 482

; SOFTWARE: Patent.pm

; SEQ ID NO 412

; LENGTH: 236

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-876-997-412

Query Match 20.6%; Score 213; DB 10; Length 236;

Best Local Similarity 32.6%; Pred. No. 4.6e-11;

Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

QY 4 GLKRKHSDELEEE---ERWEWSAGLQSYQQA-----LLRISLDKVQSRSLGPRAPS 51

Db 5 GLKRKREEEKEPLAVDSW-WLDPGHTAVAQAPPAVASSSLFDLSVLKLHHSLOQSEPD 63

QY 52 LRRHVLHNTLQQLQAALRLAPALPPEPL-----FLGEEDFSLSATIGSILRE 101

Db 64 LRHLVLVXNTRLRIQAS--MAPAAALPPVTPPAAPXVADNLLASSDAALSASMAXLLED 121

QY 102 LDTSMGTEPPQNPTPLGLQNEVPPD-----QPDVFLEAL--SSRYLGDSGLDDFFL 152

Db 122 L-SHIEGLSQAPQ-----LADEGPPGRSIGGAAPSLGALDLLGPATGCLLDNGLEGLFE 175

QY 153 DIDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEIILGS 196

Db 176 DIDTSMYDNELWAPASEGLKPGPED---GPGKEEAPELDEAELDYLMDVLVGT 225

RESULT 7

US-10-296-115-1065

; Sequence 1065, Application US/10296115

; Publication No. US20040053248A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq Inc

; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides

; FILE REFERENCE: 784PCT

; CURRENT APPLICATION NUMBER: US/10/296,115

; CURRENT FILING DATE: 2002-11-18

; PRIOR APPLICATION NUMBER: US09/488,725

; PRIOR FILING DATE: 2000-01-21

; PRIOR APPLICATION NUMBER: US09/552,317

; PRIOR FILING DATE: 2000-04-25

; NUMBER OF SEQ ID NOS: 1478

; SEQ ID NO 1065

; LENGTH: 278

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-296-115-1065

Query Match 20.6%; Score 213; DB 15; Length 278;

Best Local Similarity 32.6%; Pred. No. 5.6e-11;

Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

QY 4 GLKRKHSDELEEE---ERWEWSAGLQSYQQA-----LLRISLDKVQSRSLGPRAPS 51

Db 47 GLKRKREEEKEPLAVDSW-WLDPGHAAVAQAPPAVASSSLFDLSVLKLHHSLOQSEPD 105

QY 52 LRRHVLHNTLQQLQAALRLAPALPPEPL-----FLGEEDFSLSATIGSILRE 101

Db 106 LRHLVLVXNTRLRIQAS--MAPAAALPPVTPPAAPXVADNLLASSDAALSASMAXLLED 163

QY 102 LDTSMGTEPPQNPTPLGLQNEVPPD-----QPDVFLEAL--SSRYLGDSGLDDFFL 152

Db 164 L-SHIEGLSQAPQ-----LADEGPPGRSIGGAAPSLGALDLLGPATGCLLDNGLEGLFE 217

QY 153 DIDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEIILGS 196

Db 218 DTDTSYDNLWAPASEGLKPGPED---GPGKEEAPELDEALDYLMVDVLVGT 267

RESULT 8

US-10-029-386-34001  
Sequence 34001, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR Q  
TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
FILE REFERENCE: AEOMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 34001  
LENGTH: 222  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC010271.5  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.46  
OTHER INFORMATION: SWISSPROT HIT: Q14140, EVALUE 2.70e-01  
US-10-029-386-34001

Query Match 19.8%; Score 205; DB 14; Length 222;  
Best Local Similarity 32.6%; Pred. No. 2.2e-10;  
Matches 75; Conservative 34; Mismatches 69; Indels 52; Gaps 13;  
Qy 4 GLKRRKHSDLSEEE---ERWESPAGLQSYQQA-----LLRISLDKQVRSGLGPRAPS 51  
Db 5 GLKRRKREEEKEPLAVDSW-WLDPGHTAVAQAAPPVAVSSSLFSLVSLKHHSLQQSEPD 63  
Qy 52 LRRHVLHNTLQQLQAALRLAPALPPEPL-----FLGEEDFSLSATIGSILRE 101  
Db 64 LRHLVLVNTLRRIQAS--MAPAAALPPVPSPPAAPSVADNLLASSDAALSASMALLED 121  
Qy 102 LDTSMGTEPPQNPVTPLGLQNEVPP-----QPDVPVLEAL--SSRYLGDSGLDDFFL 152  
Db 122 L-SHIEGLSQAPQP-----LADEGPPGRSIGGAAPSLGALDILGPATGCLLDDGLEGLPE 175  
Qy 153 DTDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEII 193  
Db 176 DTDTSYDNLWAPASEGLKPGPED---GPGKEEAPELDEALDYLMVDVL 222

RESULT 9

US-10-076-069-2  
Sequence 2, Application US/10076069  
Publication No. US20020177214A1  
GENERAL INFORMATION:  
APPLICANT: JURECIC, ROLAND  
APPLICANT: NACHTMAN, RONALD  
TITLE OF INVENTION: HEPP, A NOVEL GENE WITH A ROLE IN HEMATOPOIETIC AND NEURAL DEVELO  
FILE REFERENCE: 39532-176599  
CURRENT APPLICATION NUMBER: US/10/076,069  
CURRENT FILING DATE: 2002-02-15  
PRIOR APPLICATION NUMBER: US 60/268,923  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 2  
LENGTH: 237  
TYPE: PRT  
ORGANISM: Mus musculus  
US-10-076-069-2

Query Match 13.2%; Score 136; DB 13; Length 237;  
Best Local Similarity 26.1%; Pred. No. 0.00043;  
Matches 63; Conservative 32; Mismatches 86; Indels 60; Gaps 9;

Qy 4 GLKRRKHSDLSEEEERWESPA-GLQSYQQALLRISLDKQVRSGLGPRAPSLRRHVLHNTL 62  
Db 5 GLKRRKYGDQEEGVEGFTVPSYSLO--RQSLDMSLVKLQLCHMLVEPNLCRSVLIANTV 62  
Qy 63 QQLQAALR-----LAP--APALPPEPLFLGEEDFSLSATIGSILRELDTSMDGTEPPQ 113  
Db 63 ROIQEEMSQDGVVHGMAPQNVDRAPVERLVSTE---ILCRTVRGAEEHPAPELEDAPLQ 119  
Qy 114 NPVTPLGLQNEVPPQPPD-----VFLEALSSRYLGDSGLDDFF 151  
Db 120 NSVSELPVIGSAPGORNPPQSSLWEMDSPQENRGSFQKSLDQIFETLENK--NSSSVEELF 177  
Qy 152 LDIDTSAVEKEPAR-----APPEPPHNLFCAPGSWEWNELDHIMEII 193  
Db 178 SDVDSSYYDLDTVLTCMMSGTKSSLCNGLGEGFAAATPPPSSTCKS---DLAELDHVVEIL 234  
Qy 194 L 194  
Db 235 V 235

RESULT 10

US-10-170-385-31  
Sequence 31, Application US/10170385  
Publication No. US20030203372A1  
GENERAL INFORMATION:  
APPLICANT: Ward, Neil Raymond  
APPLICANT: Mundy, Christopher Robert  
APPLICANT: Kan, On  
APPLICANT: Harris, Robert Alan  
APPLICANT: White, Jonathan  
APPLICANT: Binley, Katie Mary  
APPLICANT: Rayner, William Nigel  
APPLICANT: Naylor, Stuart  
APPLICANT: Kingsman, Susan Mary  
APPLICANT: Krige, David  
TITLE OF INVENTION: ANALYSIS METHOD  
FILE REFERENCE: 532682000100  
CURRENT APPLICATION NUMBER: US/10/170,385  
CURRENT FILING DATE: 2002-06-12  
PRIOR APPLICATION NUMBER: PCT/GB02/01662  
PRIOR FILING DATE: 2002-04-08  
PRIOR APPLICATION NUMBER: PCT/GB01/05458  
NUMBER OF SEQ ID NOS: 549  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 31  
LENGTH: 314  
TYPE: PRT  
ORGANISM: Homo Sapiens  
US-10-170-385-31

Query Match 12.1%; Score 125; DB 14; Length 314;  
Best Local Similarity 22.7%; Pred. No. 0.006;  
Matches 71; Conservative 34; Mismatches 88; Indels 120; Gaps 12;  
Qy 2 EGGLKRKHSDLSEEEERWESPAQLQS-----YQALLRISLDKQVRSGLGPRAPSLRRH 55  
Db 4 KGG-KRKFDHEDDGLGKIVSPCDGFSKVSYTLQRQTIFNLSMLKLYNHRPLTEPSLQKT 62  
Qy 56 VLIHNTLQQLQAALR-----PALPPE---PLFLGHEDFSLSATIGSILRELD 103  
Db 63 VLINMLRRIQEELKQEGSLRPMFTPSQPTTEPSDSYREAPPAPASHLASPSHPCDLGS 122

Qy 71 -----LAPA-----PALPPE---PLFLGHEDFSLSATIGSILRELD 103  
Db 123 TTPLEACLTPTASLLEDDDDTFTCTSQAMQPTAPTKLSPALLPEKD-SFSSALDEIEELCP 181  
Qy 104 TSMD-----GTEPPQNPVTPLGLQNEVPPQ-----PDPVFLEAL-----SSRYLGDS 145  
Db 182 TSTSTEATAATADSVKGTSSSEAGTQKLDGQPQESRADDSKLMDSLPGNFEITTTGTGFLTDL 241







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2004, 13:58:11 ; Search time 40 Seconds  
(without alignments)  
324.958 Million cell updates/sec

Title: US-10-069-386-2  
Perfect score: 1033  
Sequence: 1 MEGGLKRXHSDLEEEERWE.....APGSWEWNELDHIMEILGS 196

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	94	9.1	433	4	US-08-259-451-3
2	93	9.0	433	1	US-07-672-483-4
3	93	9.0	2618	3	US-09-413-814-28
4	92.5	9.0	432	5	PCT-US95-04910-13
5	90.5	8.8	370	4	US-09-377-285B-16
6	90.5	8.8	947	2	US-08-887-518-2
7	90.5	8.8	947	2	US-09-023-321-2
8	90.5	8.8	947	2	US-09-032-475-2
9	90.5	8.8	947	3	US-09-257-703-1
10	90.5	8.8	947	4	US-09-871-889A-1
11	89.5	8.7	215	4	US-08-778-717-9
12	88.5	8.6	897	4	US-09-849-602-18
13	88	8.5	1130	4	US-09-976-594-280
14	87.5	8.5	709	4	US-10-118-328-4
15	86.5	8.4	136	4	US-08-259-451-5
16	84	8.1	337	4	US-09-543-681A-7444
17	84	8.1	514	4	US-09-252-991A-25281
18	84	8.1	628	4	US-09-345-473E-48
19	84	8.1	2152	3	US-09-036-987A-3
20	84	8.1	2152	3	US-09-370-700-3
21	84	8.1	2152	4	US-09-603-207-3
22	83	8.0	969	2	US-08-548-159-1
23	83	8.0	986	2	US-08-548-159-3
24	83	8.0	1012	3	US-08-811-481-16
25	83	8.0	1012	4	US-09-876-527-16
26	82.5	8.0	1130	4	US-09-538-092-834
27	82.5	8.0	2142	4	US-09-538-092-1142

28	82	7.9	667	4	US-09-248-796A-18663	Sequence 18663, A
29	82	7.9	1006	4	US-09-023-905A-12	Sequence 12, Appl
30	82	7.9	2101	1	US-08-466-390-4	Sequence 4, Appli
31	82	7.9	2101	1	US-08-470-950-4	Sequence 4, Appli
32	82	7.9	2101	1	US-08-467-781-4	Sequence 4, Appli
33	82	7.9	2101	1	US-08-195-487-4	Sequence 4, Appli
34	82	7.9	2101	2	US-08-483-924-4	Sequence 4, Appli
35	82	7.9	2101	3	US-09-452-294-1	Sequence 1, Appli
36	82	7.9	2101	5	PCT-US93-06160-4	Sequence 4, Appli
37	81.5	7.9	1597	3	US-09-423-890-13	Sequence 13, Appl
38	81.5	7.9	1597	3	US-08-628-829-14	Sequence 14, Appl
39	81	7.8	357	4	US-09-252-991A-26979	Sequence 26979, A
40	81	7.8	535	3	US-08-813-574-2	Sequence 2, Appli
41	80.5	7.8	513	4	US-09-732-025-2	Sequence 2, Appli
42	80.5	7.8	583	4	US-09-252-991A-32735	Sequence 32735, A
43	80.5	7.8	804	4	US-09-270-767-46750	Sequence 46750, A
44	80.5	7.8	1162	2	US-08-728-323A-2	Sequence 2, Appli
45	80.5	7.8	1162	3	US-09-298-568-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-08-259-451-3  
; Sequence 3, Application US/08259451  
; Patent No. 6406841  
; GENERAL INFORMATION:  
; APPLICANT: Lee, Helen H.  
; APPLICANT: Swanson, Priscilla A.  
; APPLICANT: Idler, Kenneth B.  
; APPLICANT: Rosenblatt, Joseph D.  
; APPLICANT: Chen, Irvin S. Y.  
; APPLICANT: Golde, David W.  
; APPLICANT: Robertson, Eugene F.  
; APPLICANT: Stephens, John E.  
; APPLICANT: Chan, Emerson W.  
; APPLICANT: Buytendorp, Mark H.  
; APPLICANT: Johnson, Joan E.  
; APPLICANT: Motley, Cheryl T.  
; APPLICANT: Peterson, Bryan.  
; APPLICANT: Edwards, Michelle  
; APPLICANT: Guidinger, Peggy  
; APPLICANT: Tate, Cynthia  
; TITLE OF INVENTION: HTLV-IINRA Compositions  
; TITLE OF INVENTION: and Assays for Detecting HTLV Infection  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Abbott Laboratories  
; STREET: One Abbott Park Road  
; CITY: Abbott Park  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60064  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/259,451  
; FILING DATE: 20-JUN-1994  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/086,415  
; FILING DATE: 01-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Daniel W. Collins  
; REGISTRATION NUMBER: 31,912  
; REFERENCE/DOCKET NUMBER: 5381.US.P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (708) 937-6365

TELEFAX: (708) 938-2623  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 433 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-259-451-3  
Query Match 9.1%; Score 94; DB 4; Length 433;  
Best Local Similarity 23.1%; Pred. No. 0.66;  
Matches 40; Conservative 24; Mismatches 67; Indels 42; Gaps 7;  
QY 47 PRAPSLRRHVLHNTLQQLQAALRLAPAPA-----LPPEPLFLGEEDFSLSA 93  
DB 13 PKAP---RGLSTHHLNLFQAAYRLQGPSDFDFQQLRRFLKALKTPILWLNPIIDYSLLA 69  
QY 94 TI-----GSILRELDTSMDGTEPPQNVPVPLG---LQNEVPPQDPVFLFA-LSSRYLG 143  
DB 70 SLIPKGYGRVVEIILVKNQVSPSAAPVPTPICPTTTTPPPPPPPSPPEAHVPPPY-- 127  
QY 144 DSGLDFFLDIDTSAVEKEPARAPPEPHNLFCAPGSWENELDHIMEIILGS 196  
DB 128 -----VEPTTQCFFPILHPPGAP-----SAHRPWQMKDLQAQKQEVSSS 166  
RESULT 2  
US-07-672-483-4  
Sequence 4, Application US/07672483  
Patent No. 5359029  
GENERAL INFORMATION:  
APPLICANT: LACROIX, Martial  
APPLICANT: ZREIN, Maan  
TITLE OF INVENTION: PEPTIDES AND ANALOGUES AND MIXTURES  
TITLE OF INVENTION: THEREOF FOR DETECTING ANTIBODIES TO HTLV-I AND HTLV-II  
TITLE OF INVENTION: VIRUSES  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 875 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10022-6250  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/672,483  
FILING DATE: 19910302  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/554,258  
FILING DATE: 18-JUL-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: HALEY Jr, James F  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: IAF8 CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)715-0600  
TELEFAX: (212)715-0674  
TELEX: 14-8367  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 433 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-672-483-4  
Query Match 9.0%; Score 93; DB 1; Length 433;  
Best Local Similarity 23.1%; Pred. No. 0.82;  
Matches 40; Conservative 24; Mismatches 67; Indels 42; Gaps 7;  
QY 47 PRAPSLRRHVLHNTLQQLQAALRLAPAPA-----LPPEPLFLGEEDFSLSA 93  
DB 13 PKAP---RGLSTHHLNLFQAAYRLQGPSDFDFQQLRRFLKALKTPILWLNPIIDYSLLA 69  
QY 94 TI-----GSILRELDTSMDGTEPPQNVPVPLG---LQNEVPPQDPVFLFA-LSSRYLG 143  
DB 70 SLIPKGYGRVVEIILVKNQVSPSAAPVPTPICPTTTTPPPPPPPSPPEAHVPPPY-- 127  
QY 144 DSGLDFFLDIDTSAVEKEPARAPPEPHNLFCAPGSWENELDHIMEIILGS 196  
DB 128 -----VEPTTQCFFPILHPPGAP-----SAHRPWQMKDLQAQKQEVSSS 166  
RESULT 3  
US-09-413-814-28  
Sequence 28, Application US/09413814  
Patent No. 6225064  
GENERAL INFORMATION:  
APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH  
APPLICANT: Bristol-Myers Squibb, Co.  
APPLICANT: Beyer, Stefan  
APPLICANT: Bloecker, Helmut  
APPLICANT: Brandt, Petra  
APPLICANT: Cino, Paul M  
APPLICANT: Dougherty, Brian A  
APPLICANT: Goldberg, Steven L  
APPLICANT: Hofle, Gerhard  
APPLICANT: Mueller, Joachim  
APPLICANT: Reichenbach, Hans  
TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or  
TITLE OF INVENTION: heteropolyketide compounds  
FILE REFERENCE: PCT/US 99/23535  
CURRENT APPLICATION NUMBER: US/09/413,814  
CURRENT FILING DATE: 1999-10-07  
EARLIER APPLICATION NUMBER: DE 198 46 493.2  
EARLIER FILING DATE: 1998-10-09  
NUMBER OF SEQ ID NOS: 107  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 28  
LENGTH: 2618  
TYPE: PRT  
ORGANISM: Sorangium cellulosum  
US-09-413-814-28  
Query Match 9.0%; Score 93; DB 3; Length 2618;  
Best Local Similarity 25.5%; Pred. No. 8.5;  
Matches 55; Conservative 18; Mismatches 79; Indels 64; Gaps 11;  
QY 31 QALLR---ISLD-----KVQSLGPRAP--SLRRHVLHNTLQQLQA 67  
DB 974 QALLRREALGDEPFFQAGNSFGLRLHAKLESFAFGKSPFITDLFQHTSIRSQAEMLSG 1033  
QY 68 ALRLAP-APALPPEPLFLGEEDFSLSATIGSILRELDTSMDGTEPPQ---NPVTPLGLQN 123  
DB 1034 SSVEAPLAGAVPQPPAAAQVASSAAKSPGERGAAATSSGLTAQPPQPHFRPIAVIGLAG 1093  
QY 124 EVPPQPD-PVFLFA-----SSRYLGDGSLD-----DFFLD 153  
DB 1094 RFPAAPODLDAFLELLTEGRCGIRFFSQAELRDEGLDANRIACHNVVPKAGFLDRADHF-D 1152  
QY 154 IDTSAVEKEPAR-APPEPHNLFCAPGSWENELDH 188  
DB 1153 ADFFGIPPRDAEITDPQIRLLECC-----WNALAH 1183  
RESULT 4  
PCT-US95-04910-13  
Sequence 13, Application PC/TUS9504910  
GENERAL INFORMATION:  
APPLICANT: The Government of the United



APPLICANT: States of America as represented  
APPLICANT: by the Secretary, Department of  
APPLICANT: Health and Human Services  
TITLE OF INVENTION: ISOLATION AND  
TITLE OF INVENTION: CHARACTERIZATION OF A NOVEL PRIMATE T-CELL  
TITLE OF INVENTION: LYMPHOTROPIC VIRUS AND THE USE OF THIS VIRUS  
TITLE OF INVENTION: OR COMPONENTS THEREOF IN DIAGNOSTIC ASSAYS  
TITLE OF INVENTION: AND VACCINES  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04910  
FILING DATE: 21-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US08/231,526  
FILING DATE: 22-APR-1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: WILLIAM S. FEILER  
REGISTRATION NUMBER: 26,728  
REFERENCE/DOCKET NUMBER: 2026-4125PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 432 amino acids  
TYPE: amino acids  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
PCT-US95-04910-13

Query Match 9.0%; Score 92.5; DB 5; Length 432;  
Best Local Similarity 21.5%; Pred. No. 0.92;  
Matches 40; Conservative 19; Mismatches 48; Indels 79; Gaps 7;  
QY 47 PRAPSLRRHVLIHNTLQQLQAALRLAPAPA-----LPPPEFLGEEDEFSLSA 93  
Db 13 PKAP---RGLSTHVLNFWLQASRYLQPGPSDFQQLRRFLKALKTPILWLPIDYSLLA 69  
QY 94 TI-----GSLRELDTSMDGTEPPQNVPVTPGLQNEVPPQDPVFLEALSSRYLGD 144  
Db 70 SLIPKGYPGRTSEINVLIRNQASPTPPAPSLP-----EPANPPPL----- 111  
QY 145 SGLDDFFLDIDTSAVEKEPARAPPEP-----PHNLFCAPGSWEWNE 185  
Db 112 -----QQPS-APPEPHTPPPVIEPPATHCLPILHPHCAPSAHRPWQMKD 154  
QY 186 LDHIME 191  
Db 155 LQAIKQ 160

RESULT 5  
US-09-377-285B-16  
Sequence 16, Application US/09377285B  
Patent No. 6720175  
GENERAL INFORMATION:  
APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
APPLICANT: WORLEY, Paul  
APPLICANT: TU, Jian

APPLICANT: XIAO, Bo  
APPLICANT: LEAHY, Daniel  
APPLICANT: BENEKEN, Jutta  
APPLICANT: LANAHAN, Anthony  
TITLE OF INVENTION: NUCLEIC ACID MOLECULE ENCODING HOMER 1b PROTEIN (AS AMENDED)  
FILE REFERENCE: JHU1580-4  
CURRENT APPLICATION NUMBER: US/09/377,285B  
CURRENT FILING DATE: 1999-08-18  
PRIOR APPLICATION NUMBER: US 60/138,426  
PRIOR FILING DATE: 1999-06-10  
PRIOR APPLICATION NUMBER: US 60/138,493  
PRIOR FILING DATE: 1999-06-10  
PRIOR APPLICATION NUMBER: US 60/138,494  
PRIOR FILING DATE: 1999-06-10  
PRIOR APPLICATION NUMBER: US 60/097,334  
PRIOR FILING DATE: 1998-08-18  
NUMBER OF SEQ ID NOS: 72  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 16  
LENGTH: 370  
TYPE: PRT  
ORGANISM: Rattus norvegicus  
US-09-377-285B-16  
Query Match 8.8%; Score 90.5; DB 4; Length 370;  
Best Local Similarity 23.5%; Pred. No. 1.2;  
Matches 43; Conservative 29; Mismatches 62; Indels 49; Gaps 8;  
QY 33 LLRISLDKV-QRSLGPRAPSLRRHVLIHNTLQQLQAALRLAPAPALPP-----EPLFLG 85  
Db 102 MVNMHLEKVARREIGTLA-----TVVRLPPSQKVIPPESLPPLTPYCRKKPL--- 147  
QY 86 EEDFSLSATIGSILRELDTSMDGT-----EPPQNPVTPGLQNEVPPQ-PDPVFLEAL-- 137  
Db 148 --NFACLDVVGHGVDLSTQLSRTGTLSRKSIAKAPATPASATLGRPPRIPEPVQLPAVDP 205  
QY 138 -----SSRYLGDSDLDDFFLDIDTSAVEKEPARAPPEP-----PHNLFCAPG 179  
Db 206 GKLSAASSVSSLASAGSAEGASGIPQSKGVAPATPPPPPIAPVTPPPPLPAEIFLLPP 265  
QY 180 SWE 182  
Db 266 PME 268

RESULT 6  
US-08-887-518-2  
Sequence 2, Application US/08887518  
Patent No. 5843721  
GENERAL INFORMATION:  
APPLICANT: Rothe, Mike  
APPLICANT: Wu, Lin  
TITLE OF INVENTION: NIK Proteins, Nucleic Acids and Methods  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/887,518  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: OSMAN, RICHARD A  
REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: T97-008  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 343-4341  
TELEFAX: (415) 343-4342  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 947 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-887-518-2

Query Match 8.8%; Score 90.5; DB 2; Length 947;  
Best Local Similarity 25.8%; Pred. No. 4;  
Matches 54; Conservative 17; Mismatches 87; Indels 51; Gaps 10;  
QY 3 GGLKRRKHSLEEEERWEWSWSPAGLQSYQQALLRISLDKQVRSGLGPRAPSLRRHVLHNTL 62  
DB 659 GGLK---SPWGEYKEPRHPPPNQANYHQ-----TLHAQPRELSPRAPGPRPAEETTGRA 710  
QY 63 QQLQAALRLAPALPPE-----PFLGEEDEFSLSATIGSILRELDTSMDGTTEPPQNPVT 117  
DB 711 PKLQ-----PPLPPEPPEPNKSPPLTSLKEE-----SGMWEPLPLSSLEPAPARNPSS 758  
QY 165 RAPPEPPHNLFCAPGSW-----EWN 184  
DB 815 KASQSSRDITLSSGVHSSWSSQAARSSSWN 843

RESULT 7  
US-09-023-321-2  
Sequence 2, Application US/09023321  
Patent No. 5844073  
GENERAL INFORMATION:  
APPLICANT: Rothe, Mike  
APPLICANT: Wu, Lin  
TITLE OF INVENTION: NIK Proteins, Nucleic Acids and Methods  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/023,321  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/887,518  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: OSMAN, RICHARD A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: T97-008  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 343-4341  
TELEFAX: (415) 343-4342  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 947 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-023-321-2  
Query Match 8.8%; Score 90.5; DB 2; Length 947;  
Best Local Similarity 25.8%; Pred. No. 4;  
Matches 54; Conservative 17; Mismatches 87; Indels 51; Gaps 10;  
QY 3 GGLKRRKHSLEEEERWEWSWSPAGLQSYQQALLRISLDKQVRSGLGPRAPSLRRHVLHNTL 62  
DB 659 GGLK---SPWGEYKEPRHPPPNQANYHQ-----TLHAQPRELSPRAPGPRPAEETTGRA 710  
QY 63 QQLQAALRLAPALPPE-----PFLGEEDEFSLSATIGSILRELDTSMDGTTEPPQNPVT 117  
DB 711 PKLQ-----PPLPPEPPEPNKSPPLTSLKEE-----SGMWEPLPLSSLEPAPARNPSS 758  
QY 118 PLGLQNEVPPQ-----PDPVFLEALSSRY-----LGDSGLDDFFLDIDTSAVEKEPA 164  
DB 759 P-ERKATVPEQELQQLQLEIEFLNLSQPPSLEEQQEILSCLSIDSLSDSDS---EKNPS 814  
QY 165 RAPPEPPHNLFCAPGSW-----EWN 184  
DB 815 KASQSSRDITLSSGVHSSWSSQAARSSSWN 843

RESULT 8  
US-09-032-475-2  
Sequence 2, Application US/09032475  
Patent No. 5854003  
GENERAL INFORMATION:  
APPLICANT: Rothe, Mike  
APPLICANT: Wu, Lin  
TITLE OF INVENTION: NIK Proteins, Nucleic Acids and Methods  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/032,475  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/887,518  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: OSMAN, RICHARD A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: T97-008  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 343-4341  
TELEFAX: (415) 343-4342  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 947 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-032-475-2

Query Match 8.8%; Score 90.5; DB 2; Length 947;  
Best Local Similarity 25.8%; Pred. No. 4;  
Matches 54; Conservative 17; Mismatches 87; Indels 51; Gaps 10;





TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 215 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: RECOMBINANT  
PUBLICATION INFORMATION:  
AUTHORS: NOBUYUKI FUJII ET AL,  
TITLE: FUSED DNA SEQUENCE, FUSED PROTEIN EXPRESSED  
TITLE: FROM SAID FUSED DNA SEQUENCE AND METHOD OF  
TITLE: EXPRESSING SAID FUSED PROTEIN  
RELEVANT RESIDUES IN SEQ ID NO: 9: FROM 1 TO 215  
US-08-778-717-9  
Query Match 8.7%; Score 89.5; DB 4; Length 215;  
Best Local Similarity 23.7%; Pred. No. 0.72;  
Matches 41; Conservative 24; Mismatches 67; Indels 41; Gaps 6;  
QY 2 EGGLKRKHSDEEBE-----ERWESWAGLQSYQQALLRLSLDKVQSLG-----46  
DB 31 EGKAQPKVEVDEDELYNCAKEMACVSAITIEEAGGSSLVPRGSEFMGQIHGSLPTP 90  
QY 47 -PRAPSLRRHVLIHNTLQQLQAALRLAPAPA-----LPPEPLFLGEEDFSL 92  
DB 91 IPKAP---RGLSTHHWLNFLQAAYRLQPRPSDFDQQLRRFLKLALKTPILWLNPIDYSLL 147  
QY 93 ATI-----GSILRELDTSMDGTTPQNPVTPLG---LQNEVPPQDPVFLEA 136  
DB 148 ASLIPKGYGRVVEIINILVKQVSPSAPAAPVTPICPTTTPPPPPPSPEA 200  
RESULT 12  
US-09-849-602-18  
Sequence 18, Application US/09849602  
Patent No. 6794501  
GENERAL INFORMATION:  
APPLICANT: Scanlan, Matthew J.  
APPLICANT: Old, Lloyd J.  
APPLICANT: Stockert, Elisabeth  
APPLICANT: Chen, Yao-Tseng  
TITLE OF INVENTION: Colon Cancer Antigen Panel  
FILE REFERENCE: L0461/7105(JRV)  
CURRENT APPLICATION NUMBER: US/09/849,602  
CURRENT FILING DATE: 2001-05-04  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 18  
LENGTH: 897  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-849-602-18  
Query Match 8.6%; Score 88.5; DB 4; Length 897;  
Best Local Similarity 24.4%; Pred. No. 5.7;  
Matches 50; Conservative 23; Mismatches 75; Indels 57; Gaps 9;  
QY 13 EEEEEERWESWAGLQSYQQ-----ALLRLSLDKVQSLGRAP-----SL 52  
DB 208 QQEAER-----QALQSLRQGGTLTGKEMSTSSIPGCLLGVALE-----GDGSPHGHASL 256  
QY 53 RRRVLIHNTLQQLQAALRLAPALPPEPLFLGEEDFSLSATIGSILRELDTSMDGTTEPP 112  
DB 257 LQHVLL---LEQARQOSTLIAVPLHGCSPPLVTGERVATSMRTVGKLPHRPLSRTQSSP- 312  
QY 113 QNPVTPGLQNEVPPQDPVFLEALSSR--YLGDSGLDDFFLDIDTSAVEKEPARAPPEP 170  
DB 313 -LPQSPQALQQLVMQQHQHQQFLEKQKQQLQLGK-----ILTKTGELPRQPTTHPEET 364  
QY 171 PHNLFCAPGSWENELDHIMEIILG 195

DB 365 E-----BELTEQQEVLLG 377  
RESULT 13  
US-09-976-594-280  
Sequence 280, Application US/09976594  
Patent No. 6673549  
GENERAL INFORMATION:  
APPLICANT: Furness, Michael  
APPLICANT: Buchbinder, Jenny  
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
FILE REFERENCE: PA-0041 US  
CURRENT APPLICATION NUMBER: US/09/976,594  
CURRENT FILING DATE: 2001-10-12  
PRIOR APPLICATION NUMBER: 60/240,409  
PRIOR FILING DATE: 2000-10-12  
NUMBER OF SEQ ID NOS: 1143  
SOFTWARE: PERL Program  
SEQ ID NO 280  
LENGTH: 1130  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: Incyte ID No. 6673549 2258794CD1  
US-09-976-594-280  
Query Match 8.5%; Score 88; DB 4; Length 1130;  
Best Local Similarity 21.8%; Pred. No. 8.6;  
Matches 52; Conservative 23; Mismatches 75; Indels 88; Gaps 11;  
QY 5 LKRKHSDEEBEERWESWAGL-----QSYQQALLRLSLD-----39  
DB 539 LKR-----LKAKEEEREAQQGFNKIWRQYKAYLK-SLDHQA VNFKQNDTKALRSKSL 593  
QY 40 -----KVQRSGLPRAPSLRRH-VLIHNTLQQLQAALRLAP-----APALPPE--- 80  
DB 594 NEIESVYDEHQEQHSEGRSAPSPHPLIFVYEDRQILEDAALISYVVKRQPAIQKEDQG 653  
QY 81 -----PLFLGEEDFSLSATIGSILRELDTSMDGTTEPPQNPVTPGLQNEVPPQPD 130  
DB 654 TIHQLLHQFVPSLFFSQQLDLGASEESADEDSDSQGTTPSERKKPAPGPHSSPPEEK 713  
QY 131 PVFLEALSSRYLGDSGLDDFFLDIDTSAVEKEPARAPPEPPH-----NLFCAPGSW 181  
DB 714 GAF-----GDA-----PATEQPP--LPPAPHPKPLDDVYSLFFANNMW 749  
RESULT 14  
US-10-118-328-4  
Sequence 4, Application US/10118328  
Patent No. 6773904  
GENERAL INFORMATION:  
APPLICANT: YAN, Chunhua et al.  
TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
FILE REFERENCE: CL001220  
CURRENT APPLICATION NUMBER: US/10/118,328  
CURRENT FILING DATE: 2002-04-09  
PRIOR APPLICATION NUMBER: 60/282,460  
PRIOR FILING DATE: 2001-04-10  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 709  
TYPE: PRT  
ORGANISM: Mus musculus  
US-10-118-328-4  
Query Match 8.5%; Score 87.5; DB 4; Length 709;  
Best Local Similarity 26.6%; Pred. No. 5.3;







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2004, 13:57:50 ; Search time 38 Seconds  
(without alignments)  
496.276 Million cell updates/sec

Title: US-10-069-386-2  
Perfect score: 1033  
Sequence: 1 MEGGLKRRKHSLEEEERWE.....APGSWEWNELDHIMEILGS 196

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101.5	9.8	728	2 H59435	phosphoinositide-3
2	97	9.4	803	2 T40514	Chaperonin hsp78p
3	97	9.4	1201	2 G86441	unknown protein [i
4	96	9.3	817	2 S51342	verprolin - yeast
5	95.5	9.2	446	2 A42029	transcription fact
6	93.5	9.1	1734	2 A54602	microtubule-associ
7	93	9.0	433	1 FOLJH2	gag polyprotein -
8	91.5	8.9	1621	2 A82255	hypothetical prote
9	90.5	8.8	804	2 AG0565	probable membrane
10	88	8.5	651	2 T42644	hypothetical prote
11	88	8.5	716	2 T26998	hypothetical prote
12	88	8.5	1676	2 A56508	anucleate primary
13	86.5	8.4	536	2 A34596	transcription fact
14	86.5	8.4	1494	2 T14355	protein-tyrosine-p
15	86	8.3	475	2 B37761	ntfA protein - Thi
16	85.5	8.3	485	2 T37550	hypothetical coile
17	85.5	8.3	825	2 E75508	conserved hypothet
18	85.5	8.3	1544	2 E59431	phosphoinositide-b
19	85	8.2	715	2 D85087	hypothetical prote
20	85	8.2	1217	2 T00270	hypothetical prote
21	84.5	8.2	366	1 SAVLWD	large surface anti
22	84.5	8.2	477	2 T46304	hypothetical prote
23	84.5	8.2	723	2 B38749	3-phosphatidylinos
24	84.5	8.2	908	2 T16057	hypothetical prote
25	84.5	8.2	982	2 T43676	hunchback-related
26	84.5	8.2	1262	2 T13353	protein stn-B - fr
27	84	8.1	289	2 D70452	leucine-tRNA ligas
28	84	8.1	954	2 I61714	co-repressor prote
29	84	8.1	1015	2 JC5062	phogrin precursor

30	84	8.1	1026	2 T28968	hypothetical prote
31	84	8.1	1234	2 T00363	hypothetical prote
32	84	8.1	1819	2 T32008	hypothetical prote
33	83.5	8.1	331	2 S09800	hypothetical prote
34	83.5	8.1	601	2 AG0066	probable AMP-bindi
35	83.5	8.1	804	2 A85549	probable oxidoredu
36	83.5	8.1	804	2 G90698	probable oxidoredu
37	83	8.0	346	2 E72672	hypothetical prote
38	83	8.0	737	2 S28030	DNA-binding protei
39	83	8.0	1353	2 T00249	ichi protein - ink
40	82.5	8.0	555	2 T30349	structural protein
41	82.5	8.0	1130	1 TVHUA	protein-tyrosine k
42	82.5	8.0	1870	2 S37671	MHC class III hist
43	82.5	8.0	1872	2 S36152	MHC class III hist
44	82.5	8.0	2142	2 B35098	MHC class III hist
45	82	7.9	450	2 B47265	tailless (tll) pro

ALIGNMENTS

RESULT 1

H59435  
phosphoinositide-3-kinase regulatory beta chain [imported] - human  
C;Species: Homo sapiens (man)  
C;Date: 03-Jun-2002 #sequence\_revision 03-Jun-2002 #text\_change 09-Jul-2004  
C;Accession: H59435; A59436  
R;Volinia S; Patracchini P; Otsu M; Hiles I; Gout I; Calzolari E; Bernardi F; Rooke L; v  
Oncogene 7, 789-793, 1992  
A;Title: Chromosomal localization of human p85 alpha, a subunit of phosphatidylinositol  
A;Reference number: H59435  
A;Accession: H59435  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-728 <VOL>  
A;Cross-references: UNIPROT:O00459; GB:NP\_005018; PID:g4826908; PIDN:NP\_005018.1  
R;Janssen, J.W.; Schleithoff, L.; Bartram, C.R.; Schulz, A.S.  
Oncogene 16, 1767-1772, 1998  
A;Title: An oncogenic fusion product of the phosphatidylinositol 3-kinase p85beta subun  
A;Reference number: A59436; MUID:98241181; PMID:9582025  
A;Accession: A59436  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-728 <JAN>  
A;Cross-references: GB:NP\_005018; PID:g4826908; PIDN:NP\_005018.1

Query Match 9.8%; Score 101.5; DB 2; Length 728;  
Best Local Similarity 23.9%; Pred. No. 1;  
Matches 47; Conservative 25; Mismatches 70; Indels 55; Gaps 9;

QY	20	EWSA----	GLQSYQQAL-----	-----LRISLDKQVRSIGPRAPSLRRHVLIH	59
		:	:	:	:
		:	:	:	:
Db	164	QWDTAALADGIKSFLALPAPLVTPPEASAEARRALREAGPVGPALEPPTLPALRALTLR	223		
QY	60	NTLQQLQAALRLAPAPALPPEPLFLGEEDFSLSATIGSILRELDTSMDGTPEPPQNPVTPL	119		
		:	:	:	:
		:	:	:	:
Db	224	FLLOHLGRVARRAPA-----	-----LGPVAVRALGATFGPLL-----	LRAPPPPPSSP-PPG	267
QY	120	GLQNEVPPQPD--PVFLEALSSRYLGDSGLDDFFLDIDTSAVEKEPARAPPEP-----	170		
		:	:	:	:
		:	:	:	:
Db	268	GAPDGEPSPDFPALLVEKLLQEHLEEQ-----	-----EVAPPALPPKPKAKPAPTVLANGG	320	
QY	171	-PHNLFCAPGSWEWNEL	186		
	:	:	:	:	:
Db	321	SPPSL--QDAEWYWGDI	335		

RESULT 2

T40514  
Chaperonin hsp78p - fission yeast (Schizosaccharomyces pombe)  
C;Species: Schizosaccharomyces pombe  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C;Accession: T40514



Db 634 ASLLHDVLPSSNLEKPPSPFVAAAPLPTFSAPSLPQQSVSTSIPTPPV-APTLSVR-- 690

QY 143 GDSGLDDFFLDIDTSAVEKEPARAPPEPP 171

Db 691 -----TETESISKNPTKSPPPPP 708

RESULT 5

A42029

transcription factor E3 - mouse

C;Species: Mus musculus (house mouse)

C;Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004

C;Accession: A42029

R;Roman, C.; Matera, A.G.; Cooper, C.; Artandi, S.; Blain, S.; Ward, D.C.; Calame, K.

Mol. Cell. Biol. 12, 817-827, 1992

A;Title: mTFE3, an X-linked transcriptional activator containing basic helix-loop-helix

A;Reference number: A42029; MUID:92123207; PMID:1732746

A;Accession: A42029

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-446 <ROM>

A;Cross-references: UNIPROT:Q64092; GB:S76673; NID:G243439; PIDN:AAB21130.1; PID:G243440

A;Note: sequence extracted from NCBI backbone (NCBIN:76673, NCBIPI:76674)

Query Match 9.2%; Score 95.5; DB 2; Length 446;

Best Local Similarity 27.3%; Pred. No. 1.7;

Matches 48; Conservative 24; Mismatches 61; Indels 43; Gaps 8;

QY 12 LEEEEERWESPAGLQSYQQALLRISLDKVORSIGLPPRAPSLRRHVLHNTLQQLQAALRL 71

Db 272 LQKEQQR---SKDLESRQR-----SLEQANRSIQLRIQEL-----ELQAQIHG 311

QY 72 APAPALPPEPLFLGEEDFSLSATIGSILRELDTSMDGTEPPQNVPVTLGLQNEVPPQDP 131

Db 312 LP--VPNPGLLSLTSSVSDSLKP--EQLDIEEGRPSTTFHVSQSPAQNAPPQQPPA 366

QY 132 VFLEAL-----SSRYLGDSGLDDFFLDIDTSAVEKE-----PARAPPEP 170

Db 367 PPSDALLDLHFPSDHGLDG-DPFHIGLEDILMEEEGMVGGLSGGALSPLRAASDP 421

RESULT 6

A54602

microtubule-associated serine/threonine protein kinase MAST205 - mouse

C;Species: Mus musculus (house mouse)

C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004

C;Accession: A54602

R;Walden, P.D.; Cowan, N.J.

Mol. Cell. Biol. 13, 7625-7635, 1993

A;Title: A novel 205-kilodalton testis-specific serine/threonine protein kinase associat

A;Reference number: A54602; MUID:94067123; PMID:8246979

A;Accession: A54602

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1734 <WAL>

A;Cross-references: UNIPROT:Q60592; GB:U02313; NID:G406057; PIDN:AAC04312.1; PID:G406058

C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

C;Keywords: ATP

F;451-726/Domain: protein kinase homology <KIN>

F;459-467/Region: protein kinase ATP-binding motif

Query Match 9.1%; Score 93.5; DB 2; Length 1734;

Best Local Similarity 26.3%; Pred. No. 15;

Matches 51; Conservative 23; Mismatches 71; Indels 49; Gaps 8;

QY 25 GLQSYQQAL-----LRISLDKVORSIGLPPRAPSLRRHVLHNTLQQLQAALRLAPALPPP 79

Db 1308 GSQSFPPTKLHLSPLGLRQLSRPXSAEPPRSPLUKR-----VQSAEKLAALAAAE 1357

QY 80 EPLFLGEEDFSLSATIGSILRELDTSMDGTBEPQNPVTLGLQNEVP---PQDPVFLEA 136

Db 1358 KKL-APSRKHSLDLPHGELKKEL-----TPREASPLEVVGTRSVLSKGPLPGKVLQP 1410

QY 137 LSSRYLG-----DSGLDDFFLDIDTSAVEKEPARAP-PEPPHN 173

Db 1411 APSRALGTLRQDRAERRESLQKQEAIREVDSSDDTDEEPNSQATQEPRLSPHEASHN 1470

QY 174 LFCAPGSWEWNELD 187

Db 1471 LL-PKSGSGEGTEED 1483

RESULT 7

FOLJH2

gag polyprotein - human T-cell lymphotropic virus type 2

N;Alternate names: core polyprotein

N;Contains: core protein p12; core protein p15; core protein p24

C;Species: human T-cell lymphotropic virus type 2, HTLV-2

A;Note: host Homo sapiens (man)

C;Date: 28-Feb-1986 #sequence\_revision 28-Feb-1986 #text\_change 09-Jul-2004

C;Accession: A03944

R;Shimotohno, K.; Takahashi, Y.; Shimizu, N.; Gojobori, T.; Golde, D.W.; Chen, I.S.Y.; Proc. Natl. Acad. Sci. U.S.A. 82, 3101-3105, 1985

A;Title: Complete nucleotide sequence of an infectious clone of human T-cell leukemia v

A;Reference number: A94042; MUID:85216449; PMID:2582407

A;Accession: A03944

A;Molecule type: DNA

A;Residues: 1-433 <SHI>

A;Cross-references: UNIPROT:P03346; GB:M10060; NID:G329559; PIDN:AAB59884.1; PID:G32956;

A;Note: the authors translated the codon TAC for residues 197 and 249 as Thr

C;Genetics:

A;Gene: gag

C;Superfamily: mammalian retrovirus gag polyprotein II

C;Keywords: core protein; polyprotein

F;1-136/Product: core protein p15 #status predicted <P15>

F;137-214/Product: core protein p24 #status predicted <P24>

F;215-433/Product: core protein p12 #status predicted <P12>

Query Match 9.0%; Score 93; DB 1; Length 433;

Best Local Similarity 23.1%; Pred. No. 2.7;

Matches 40; Conservative 24; Mismatches 67; Indels 42; Gaps 7;

QY 47 PRAPSLRRHVLHNTLQQLQAALRLAPAPA-----LPPEPLFLGEEDFSLSA 93

Db 13 PKAP---RGLSTHHWLNFLQAAYRLQRPSPDFDFOQLRRFLKLALKTPIWLNPIDYSLLA 69

QY 94 TI-----GSILRELDTSMDGTEPPQNVPVTLG--LQNEVPPQDPVFLEA-LSSRYLG 143

Db 70 SLIPKGYPGRVVEIINILVKNVQSPSAPAAPVPFICPTTTPPPPPPSPEAHVPPPY-- 127

QY 144 DSGLDDFFLDIDTSAVEKEPARAPPEPPHNLFCAPGSWEWNELDHIMEILGS 196

Db 128 -----VEPTTTCQFPILHPPGAP----SAHRPWQMKDLQAIKQEVSSS 166

RESULT 8

A82255

hypothetical protein VC0998 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C;Species: Vibrio cholerae

C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004

C;Accession: A82255

R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, I l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: A82255

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1621 <HEI>

A;Cross-references: UNIPROT:Q9KTA5; GB:AE004181; GB:AE003852; NID:G9655454; PIDN:AAF941E

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC0998



A;Map position: 1

Query Match 8.9%; Score 91.5; DB 2; Length 1621;  
Best Local Similarity 27.4%; Pred. No. 20;  
Matches 49; Conservative 22; Mismatches 63; Indels 45; Gaps 8;

QY 13 EEEERWESPAGLQSYQQALLRISLDKQVRSGLGRAPSLRRHVLHNTLQQLQAALRLA 72

DB 530 EEDDEFDLSCAGVAGDQ-----DLDDLFAIEEQA-----DLEQLEA--KAI 570

QY 73 PAPALPPEPLFLGEEDEFSLSATIGSILRELDTSMGTE-----PPQNVPVTEL- 119

DB 571 DETALLDE--ILAEQDAPLSEESTELLDELDDDFDKFENDEFDAQTADLLQPEEPILDLE 628

QY 120 ----GLQNEVPQPDVPVFLA-----LSSRYLGDGLDDFFLDIDTSAVEKEPARAPPE 169

DB 629 EDSTQLLNEVLGEPVPEELASGLEIDQNSTELLDELDD--LDLDDESIATEFSVAPE 685

RESULT 9

AG0565

probable membrane protein STY0554 [imported] - Salmonella enterica subsp. enterica serov

C;Species: Salmonella enterica subsp. enterica serovar Typhi

A;Note: This species has also been called Salmonella typhi

C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002

C;Accession: AG0565

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A;Reference number: AB0502; MUID:21534947; PMID:11677608

A;Accession: AG0565

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-804 <PAR>

A;Cross-references: GB:AL513382; PIDN:CAD04993.1; PID:g16501778; GSPDB:GN00176

C;Genetics:

A;Gene: STY0554

C;Superfamily: Escherichia coli probable membrane protein ybbp

Query Match 8.8%; Score 90.5; DB 2; Length 804;

Best Local Similarity 28.7%; Pred. No. 9.9;

Matches 37; Conservative 16; Mismatches 39; Indels 37; Gaps 7;

QY 38 LDKVQVRSGLGRAPS-----LRRHVLHNTLQQLQAALRLAPALPPE 80

DB 478 LDRWQQQLPPESPYNFLINIASEQVAPLKAFLAEHQVIPQTFYPIVRA-RLTEINGNPT 536

QY 81 PLFLGEEDEFSLSATIGSILRELDTSMGTEBPQNVPVTLGLQNEVPQPDVPFL-EALSS 139

DB 537 ----GQODESLN-----RELNLTWQDTRPAHNPL----VAGHWDPKPGEVSMEEGLAK 581

QY 140 RY---LGDS 145

DB 582 RLNVKLGDS 590

RESULT 10

T42644

hypothetical protein DKFZp566N1047.1 - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jul-2004

C;Accession: T42644

R;Ottenwaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, November 1999

A;Reference number: Z2231

A;Accession: T42644

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-651 <AAA>

A;Cross-references: UNIPROT:Q9UF93; EMBL:AL133109

A;Experimental source: fetal kidney; clone DKFZp566N1047

C;Genetics:

A;Note: DKFZp566N1047.1

Query Match 8.5%; Score 88; DB 2; Length 651;

Best Local Similarity 24.4%; Pred. No. 12;

Matches 38; Conservative 26; Mismatches 68; Indels 24; Gaps 7;

QY 27 QSYQQALLRISLDKQVRSGLGRAPSLRRHVLHNTLQQLQAALRLA--PAPALPPEPL-- 82

DB 320 QSQQQPPQQQSPQPPQPPQPM-AGPLVTQSVQGLQASSQSVQYPAVSFPQHLLP 378

QY 83 -----FLGEEDEFSLSATIGSILRELDTSMGTEBPQNVPVTLGLQNEVPQP--DPVFL 134

DB 379 VSPTQHFPMRDD--VATQFGQMTLSRQSSGETPEPPSGPVYPSLM----PQPAQQPSYV 432

QY 135 EALSSRYLGDGLDDFFLDIDTSAVEKEPARAPPEP 170

DB 433 IASTGQQLPTGFGSG-----SGPPISQQLVQPPSP 463

RESULT 11

T26998

hypothetical protein Y48B6A.6 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C;Accession: T26998

R;Wall, M.

submitted to the EMBL Data Library, September 1999

A;Reference number: Z20297

A;Accession: T26998

A;Status: preliminary; translated from GB/EMBL/DBBJ

A;Molecule type: DNA

A;Residues: 1-716 <WIL>

A;Cross-references: UNIPROT:Q9U2A6; EMBL:AL110490; NID:el542263; PIDN:CAB54442.1; CESP:Y

A;Experimental source: clone Y48B6A

C;Genetics:

A;Gene: CESP:Y48B6A.6

A;Introns: 38/3; 196/3; 437/2; 460/3; 518/1; 617/3; 673/3

Query Match 8.5%; Score 88; DB 2; Length 716;

Best Local Similarity 25.0%; Pred. No. 14;

Matches 49; Conservative 29; Mismatches 80; Indels 38; Gaps 9;

QY 7 RKHSDLEEEERWEWS--PAGLQSYQQALLRISLDKQV--RSLGRAPSLRRHVLHNT 61

DB 219 RSDESEIEEERRRKESEETASFEELAEIMRISRSPVPPVLSIPPPPP--NIPPLPTI 275

QY 62 LQQLQAALRLAPALPPEPLFLGEEDEFSLSATIGSI-----LRELD 103

DB 276 PQEVQSPSPRPTSVPPPIPSGPGSEDMDELIESFSDSVIFNNSMSPPLPLPRE-- 333

QY 104 TSMGTE-PPQNVPVTLGLQNEVPQPDVPVLEALS-SRYLGDGLDDFFLDIDTSAVEK 161

DB 334 SSLETLEVTPEDPVTEKXV--EASPTPLPKATESLNESSIKALEGLEVKALEAQEASDDR 391

QY 162 EPARAP-----PEPP 171

DB 392 PSAPTPIRDSLLPPPP 407

RESULT 12

A56508

anucleate primary sterigmata A (apsA) protein - Emericella nidulans

C;Species: Emericella nidulans, Aspergillus nidulans

C;Date: 05-Jan-1996 #sequence\_revision 05-Jan-1996 #text\_change 09-Jul-2004

C;Accession: A56508

R;Fischer, R.; Timberlake, W.E.

J. Cell Biol. 128, 485-498, 1995

A;Title: Aspergillus nidulans apsA (anucleate primary sterigmata) encodes a coiled-coil

A;Reference number: A56508; MUID:95164553; PMID:7860626

A;Accession: A56508

A;Status: preliminary

A;Molecule type: DNA  
A;Residues: 1-1676 <FIS>  
A;Cross-references: UNIPROT:Q00083; GB:X82289; NID:G683499; PID:G683500  
C;Genetics:  
A;Gene: apsA  
A;Introns: 149/3

RESULT 13  
A34596  
transcription factor E3 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 06-Jul-1990 #sequence\_revision 09-Oct-1992 #text\_change 09-Jul-2004  
C:Accession: A34596; S10379  
R:Beckmann, H.; Su, L.K.; Kadesch, T.  
Genes Dev. 4, 167-179, 1990  
A:Title: TFE3: a helix-loop-helix protein that activates transcription through the immunoglobulin enhancer  
A:Reference number: A34596; MUID:90249724; PMID:2338243  
A:Accession: A34596  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-536 <BEC>  
A:Cross-references: UNIPROT:P19532; EMBL:X51330; NID:g37061; PIDN:CAA35714.1; PID:g133533  
A>Note: the authors translated the codon ACC for residue 433 as Ser, and GAG for residue 434  
C:Genetics:  
A:Gene: GDB:TFE3  
A:Cross-references: GDB:125870; OMIM:314310  
A:Map position: Xp11.23-Xp11.22  
C:Keywords: DNA binding; transcription factor

RESULT 14  
T14355  
protein-tyrosine-phosphatase (EC 3.1.3.48) TD14 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: T14355

R;Cao, L.; Zhang, L.; Ruiz-Lozano, P.; Yang, Q.; Chien, K.R.; Graham, R.M.; Zhou, M. J. Biol. Chem. 273, 21077-21083, 1998

A;Title: A novel putative protein-tyrosine phosphatase contains a BRL1-like domain and s

A;Reference number: Z18004; MUID:98361981; PMID:9694860

A;Accession: T14355

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-1494 <CAO>

A;Cross-references: UNIPROT:O88902; EMBL:AF077000; NID:g3598973; PIDN:AAC6

A;Experimental source: brain

C;Genetics:

A;Gene: PTP-TD14

C;Function:

A;Description: may be involved in regulating Ha-ras-dependent cell growth

C;Keywords: phosphoric monoester hydrolase

RESULT 15  
B37761  
ntrA protein - Thiobacillus ferrooxidans  
C;Species: Thiobacillus ferrooxidans  
C;Date: 31-May-1991 #sequence\_revision 31-May-1991 #text\_change 09-Jul-2004  
C;Accession: B37761  
R;Berger, D.K.; Woods, D.R.; Rawlings, D.E.  
J. Bacteriol. 172, 4399-4406, 1990  
A;Title: Complementation of Escherichia coli sigma(54) (ntrA)-dependent formate hydrogen  
A;Reference number: A37761; MUID:90330545; PMID:2198257  
A;Accession: B37761  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-475 <BER>  
A;Cross-references: UNIPROT:P24695; GB:M58480; GB:M33831; NID:G154642; PIDN:AAA27379.1;  
C;Superfamily: Pseudomonas transcription initiation factor sigma  
C;Keywords: DNA binding; transcription regulation

Search completed: November 15, 2004, 14:07:21

us-10-069-386-2.rpr

Mon Nov 15 17:33:07 2004

Job time : 40 secs